Monte Carlo studies of the GCT for the Cherenkov Telescope Array and the search for VHE AGN using cluster analysis

ARMSTRONG, THOMAS,PATRICK

Monte Carlo studies of the GCT for the Cherenkov Telescope Array and the search for VHE AGN using cluster analysis

Mr Thomas Patrick Armstrong

A thesis presented for the degree of
Doctor of Philosophy

Very High Energy Gamma-Ray Group
Department of Physics
University of Durham
United Kingdom

June 2016

Monte Carlo studies of the GCT for the Cherenkov Telescope Array and the search for VHE AGN using cluster analysis

Thomas Patrick Armstrong

Submitted for the degree of Doctor of Philosophy
June 2016

Abstract

The future ground based gamma-ray observatory, the Cherenkov Telescope Array (CTA), will soon enter its construction phase. This work therefore looks into providing a better model of one of the small size telescopes, the Gamma Cherenkov Telescope (GCT), for input into Monte Carlo simulations. Evaluation of these models shows that both a telescope equipped with MaPM (GCTM) and SiPM (GCTS) detector modules should meet if not exceed certain CTA requirements. To determine possible early science deliverables, a study into the performance of a small 7 telescope array, along with an extrapolation up to the full complement of SSTs is presented. This reveals promising results for both configurations, with GCTS performing better than GCTM. This work also presents an investigation into the use of local muons as a form of absolute calibration of the GCT telescope. It has been shown that, while there are some difficulties, the method should be possible. The remainder of this thesis presents results obtained from applying the clustering algorithm DBSCAN to the *Fermi*-LAT data in the very high energy (VHE) regime. This includes 9 sources detected in the Pass 7 reprocessed data set and 70 in the improved Pass 8 data set. These sources represent promising candidates for follow-up observations with current ground-based gamma-ray observatories and helps to frame the science goals of CTA.

Declaration

The work in this thesis is based on research carried out in the Very High Energy Gamma Ray Group, the Department of Physics, Durham. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Acknowledgements

I would like to start by thanking my supervisors and colleagues in Durham, Prof. Paula Chadwick, Dr. Sam Nolan and Dr. Anthony Brown. They have dedicated plenty of time and patience helping me prepare, reading drafts and answering many questions. As a member of GCT and several working groups within CTA, I have also received support and advice from countless others, many I consider friends as well as colleagues. To my family, my friends here in Durham and especially Cass, I say thanks for the patience, understanding and for keeping me sane by reminding me there is always a life outside of work.

Chapter Header Images

At the start of each chapter there is a header image, for aesthetic reasons the captions are placed here

Introduction: *The first proposed array of IACTs, a configuration that was later adopted for VERITAS (NASA Workshop, Space Lab. Science, Baton Rouge, 1984, taken from talk by Trevor Weeks: "Very High Energy Gamma Ray Astronomy 101")*

Chapter 1: *Colour composite image of Centuras A, highlighting the scale of the relativistic jets present in some galaxies. Credit: ESO/WFI (Optical); MPIfR/ESO/APEX/A.Weiss et al. (Submillimetre, orange); NASA/CXC/CfA/R.Kraft et al. (X-ray, blue)*

Chapter 2: *Front view of the GCT prototype telescope showing the back of the secondary mirror and the petals of the prototype primary mirror layout, photo credit: Harm Schoorlemmer*

Chapter 3: *Camera focal plane, combining simulation (left) and real image of the MaPM modules in the first camera prototype (right).*

Chapter 4: *Simulated camera image of a muon ring for GCTS, showing only pixels within $\pm 0.26^\circ$ of the fitted muon ring.*

Chapter 5: DBSCAN *algorithm applied to sampled data of the Durham University logo including random added noise*

Chapter 6: *Example counts map of brightest source found with* DBSCAN

Chapter 7: *Artists impression of CTA. Credit: G. Prez, IAC (SMM)*

Contents

Abstract ii

Declaration iii

Acknowledgements iv

Chapter Header Images v

1 Introduction 1

2 The Violent Universe 4

2.1 Cosmic Rays and Gamma-Rays 5

2.2 Interaction and Production Processes 6

2.2.1 Electromagnetic 8

2.2.2 Hadronic 9

2.3 Detection and Attenuation Processes 11

2.3.1 Interactions With Light 11

2.3.2 Interaction with Matter 12

2.4 The Extragalactic Sky and Active Galactic Nuclei 13

2.4.1 VHE Blazars 16

2.4.2 Relativistic Jets 16

2.4.3 Leptonic I: Synchrotron Self Compton (SSC) 19

2.4.4 Leptonic II: External Compton Radiation (ECR) 19

2.4.5 Hadronic 19

2.4.6 The Path to Understanding 20

3 Observational Techniques and the Future of γ-ray Astronomy 23

3.1 Introduction . . . 24
3.2 Space Based: Fermi . . . 24
3.3 Ground Based: Imaging Atmospheric Cherenkov Telescopes . . . 28
3.3.1 Electromagnetic Air Showers . . . 28
3.3.2 Hadronic Air Showers . . . 29
3.3.3 Cherenkov Radiation . . . 32
3.3.4 Shower Imaging . . . 35
3.4 Cherenkov Telescope Array . . . 40
3.5 GCT: High energy SST component of CTA . . . 44
3.5.1 Structure: SST-GATE . . . 45
3.5.2 Camera: CHEC . . . 48
3.5.3 Prototype . . . 52

4 Monte Carlo Simulations for GCT and CTA 55

4.1 Introduction . . . 56
4.2 Simulating Air Showers: CORSIKA . . . 57
4.3 Optics and Electronics: SIM_TELARRAY . . . 59
4.4 Simple Analysis and Processing, READ_CTA . . . 60
4.5 Advanced Analysis, TMVA . . . 60
4.6 Building GCT . . . 63
4.6.1 Optics . . . 63
4.6.2 Detectors . . . 67
4.6.3 Electronics . . . 73
4.6.4 Night Sky Background . . . 76
4.6.5 Discriminator Threshold . . . 78
4.6.6 Summary . . . 78
4.7 Low Level Evaluation . . . 80
4.7.1 Image Amplitude Trigger Efficiency . . . 80
4.7.2 Charge Resolution . . . 81
4.8 High Level Evaluation: Mini Array . . . 83
4.8.1 MC SST Mini-Production . . . 83

4.8.2 Reconstruction and Initial Cuts 84
4.8.3 Angular Resolution 85
4.8.4 Energy Resolution 87
4.8.5 Effective Area 89
4.8.6 Energy Threshold 89
4.8.7 Sensitivity 92
4.9 High Level Evaluation: CTA Sub-Array 96
4.10 Conclusion 108

5 Muon Calibration for GCT **109**
5.1 Introduction 110
5.2 Muons 111
5.3 Motivation for Calibration 112
5.3.1 Monitoring Optical Efficiency 113
5.3.2 Monitoring the Point Spread Function 114
5.3.3 Monitoring Flat Fielding 114
5.4 Muon Spectrum 114
5.5 Muon Air Shower Geometry 118
5.6 Determining Muon Efficiency 121
5.7 Systematic Effects 123
5.7.1 Instrumentation Effects 124
5.7.2 Atmospheric Effects 126
5.8 Evaluation of Method with MC Simulations 127
5.8.1 Pre-Selection Cuts 128
5.8.2 Energy Reconstruction, Circle Fitting 129
5.8.3 Post Fit Cuts 129
5.8.4 Selection Efficiency Post Cuts 131
5.8.5 Ring Broadening, Fitting *Arcwidth* 131
5.8.6 Reconstructing Impact from Modulation 135
5.9 Expected Rates and Summary Results 143
5.10 Monitoring System Degradation 144
5.11 Conclusion 147

6 Source Detection in Sparse Data Sets: Initial Study 149

6.1 Introduction . . . 150

6.2 Clustering Algorithms . . . 152

6.2.1 DBSCAN . . . 153

6.2.2 Clustering of VHE Gamma-Ray Events . . . 157

6.3 Verification of VHE Clusters using *Fermi* Analysis . . . 162

6.4 Discussion . . . 165

6.4.1 DBSCAN Performance . . . 167

6.4.2 Detected VHE Sources . . . 169

6.5 Conclusion . . . 170

7 Source Detection in Sparse Data Sets: Pass 8 173

7.1 Introduction . . . 174

7.2 Improvements to DBSCAN . . . 176

7.2.1 Variable DBSCAN Input . . . 176

7.2.2 Poisson significance estimation . . . 178

7.3 Background Extrapolation . . . 179

7.4 Data . . . 183

7.5 Maximum Likelihood Follow-up . . . 183

7.6 100 GeV - 3 TeV Cluster Results . . . 184

7.6.1 Revisiting the Probability Cut . . . 185

7.6.2 E <100 GeV and Variability . . . 188

7.6.3 Unknown sources and possible associations . . . 189

7.7 Comparison with 2FHL . . . 194

7.7.1 Cluster results . . . 195

7.7.2 Post 2FHL . . . 198

7.8 Discussion and Conclusion . . . 200

8 Summary, Conclusion and Future Work 202

A List of Acronyms and Abbreviations 219

B Full List of sim_telarray Configuration Parameters 221

List of Figures

2.1 Gamma-rays and the Electromagnetic Spectrum 6
2.2 Gamma-ray Production Mechanisms 7
2.3 Gamma-ray Interaction Mechanisms 11
2.4 Unified AGN Geometry 13
2.5 Known High Energy AGN 15
2.6 Example AGN SED 18
2.7 CTA Extragalactic Survey 21

3.1 Diagram of the *Fermi* LAT Tracker 25
3.2 Fermi Pass 7 Reprocessed and Pass 8 PSF and Energy Resolution . . 27
3.3 Electromagnetic EAS 30
3.4 Hadronic EAS 31
3.5 Cherenkov Radiation 33
3.6 Cherenkov Radiation as a Function of Altitude 34
3.7 XY Shower Profile 35
3.8 Stereoscopic Reconstruction of EAS 36
3.9 Hillas Parametrisation 37
3.10 CTA Concept 40
3.11 Example CTA Layouts 42
3.12 CTA South Expected Performance 43
3.13 GCT Structure (SST-GATE) 46
3.14 Final and Prototype Mirror Designs 46
3.15 Internal Structure of the CHEC Camera 49
3.16 GCT Camera (CHEC) Mechanics 50

3.17 GCT Prototype in Paris 52
3.18 One of the First Shower Image Recorded by GCT 53

4.1 IACT CORSIKA Package 58
4.2 TMVA Training Parameters 61
4.3 TMVA ζ Cut 62
4.4 Mirror Positions 64
4.5 Mirror Reflectivity 64
4.6 Cherenkov Light Compared to NSB 66
4.7 Raytracing of H.E.S.S. with ROBAST 68
4.8 Shadowing 69
4.9 Window Angular and Wavelength Response 70
4.10 Photon Detection Efficiency 71
4.11 Single Photo-electron Response 72
4.12 Discriminator/FADC Pulse Shape 74
4.13 Super Pixel Trigger 74
4.14 Convolved NSB Spectrum 76
4.15 Discriminator Threshold Determination 77
4.16 Configuration Summary 79
4.17 Image Size Trigger Efficiency 81
4.18 Charge Resolution 82
4.19 MC SST Mini-Production Array Layout 83
4.20 Mini-Array Angular Resolution 86
4.21 Example Fit for Energy Resolution 87
4.22 Mini-Array Energy Resolution and Bias 88
4.23 Mini-Array Effective Area 90
4.24 Mini-Array Energy Threshold 91
4.25 Mini-Array Differential Sensitivity 93
4.26 ON/OFF Background Estimation 94
4.27 Proton Rate Extrapolation 95
4.28 Triggered Events True Core Position, Part 1 96
4.29 Triggered Events True Core Position, Part 2 97

4.30 Number of Images in Reconstruction 98
4.31 CTA Subarray Angular Resolution 100
4.32 CTA Subarray Energy Resolution and Bias 101
4.33 Extrapolated Array 102
4.34 CTA Subarray Effective Area 103
4.35 CTA Subarray Differential Sensitivity 104
4.36 Direct GCTM and GCTS Comparison 105

5.1 Muon Flux as a Function of Altitude 116
5.2 Imaging Muon Rings 119
5.3 Muon Reconstruction Parameters 120
5.4 GCT Optical Efficiency to Muon and Gamma showers 124
5.5 Image Total p.e. and Pixels Pre and Post First Cut 128
5.6 Reconstructed Muon Radius 130
5.7 Effect of Camera Edge Cut 131
5.8 Muon Selection Efficiency for Each Cut 132
5.9 Determining *Arcwidth* of the Muon Ring 133
5.10 Arcwidth Compared to GCT PSF 134
5.11 Broadening of Muon Rings 135
5.12 Modelling Modulation of Light Along The Muon Ring 136
5.13 Ring Modulation Fit Example 140
5.14 Reconstructed Impact Distance 140
5.15 Reconstructed Muon Efficiency and Predicted Image Size 141
5.16 Muon Efficiency Reconstruction Bias 142
5.17 Selection Efficiency For Increasing Optical Degradation 145
5.18 Muon Rate and Efficiency as a Function of Optical Efficiency 145
5.19 Muon Efficiency Reconstruction Bias 146

6.1 DBSCAN Algorithm 155
6.2 Fermi LAT Pass 7 PSF 156
6.3 Distribution of Extragalactic Events >100 GeV 158
6.4 DBSCAN Clusters Found 160

6.5 DBSCAN Clusters Found Spatially Coincident with TeVCat Sources . 160
6.6 Unbinned Analysis Summary 162
6.7 Example TS Maps 164
6.8 DBSCAN Clusters Found to be Significant (TS >25) 165
6.9 Source Purity and Comparison of TS and Likelihood Ratio Statistic . 168
6.10 Distribution of Spectral Index Compared to Known AGN 171

7.1 Variable *MinPts* Contour Map 175
7.2 *Fermi*-LAT Pass 8 PSF 177
7.3 Powerlaw Fit to the *Fermi* Isotropic Background 179
7.4 Distribution of Fit Quality for Galactic Background 180
7.5 Spatial Distribution of Fit Quality for Galactic Background 181
7.6 Significant Clusters in 100 GeV to 3 TeV Range 184
7.7 Distribution of TS Values for all Clusters 185
7.8 Distribution of Time Between Events 189
7.9 Associations of Unkn. Sources (Part 1) 191
7.10 Associations of Unkn. Sources (Part 2) 192
7.11 A Comparison of Clusters Found with the 2FHL 195
7.12 Spectral Index and Flux of 2FHL Sources: Found Vs. Missed 196
7.13 Cluster Results Complementing 2FHL 197
7.14 Cluster Results 2FHL With Aditional 6 Months 199

List of Tables

4.1 Polynomial Mirror Shape 65
4.2 Mini Array Data, Selected Inputs 84
4.3 GCTM Subarray Summary Results 106
4.4 GCTS Subarray Summary Results 107

5.1 GCT Muon Calibration Results 143

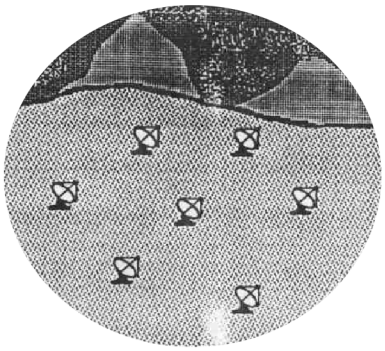
6.1 *Fermi*-LAT Data Preparation 159
6.2 DBSCAN Results Table: TeVCat Sources 161
6.3 DBSCAN Results Table 166

7.1 Cluster Results for 100 GeV to 3 TeV Range (Part 1) 186
7.2 Cluster Results for 100 GeV to 3 TeV Range (Part 2) 187
7.3 Results Table 2FHL Match 198
7.4 Results Table 2FHL Range + 6 Months 199

B.1 Configuration file parameters part 1 222
B.2 Configuration file parameters part 2 223
B.3 Configuration file parameters part 3 224

Chapter 1

Introduction



In the intrepid search for an understanding of our place within the universe, large steps are often necessary. In the coming years a new observatory will provide such a step, opening up our view of the universe to the highest energies. The Cherenkov Telescope Array (CTA) is a multi-national project, bringing knowledge from the 32 involved countries; it will provide order-of-magnitude improvements over current experiments operating in the energy range of several tens of GeV to many hundreds of TeV. Operating as an open observatory and covering both hemispheres, each array will consist of a large number of telescopes ($\approx$ 100 in the south with a smaller array in the north) and in order to cover the large energy range there will be three different sizes of telescopes. For the highest energy observations, the optimal configuration is a large number of small size telescopes (SST). In this work, it is one of the proposed solutions to these SSTs that will be considered, the Gamma-ray Cherenkov Telescope (GCT).

In the early period of a new project, there are always a wide range of jobs to

do. Therefore the work in this thesis can be grouped into three main themes: 1) Configuring and evaluating the GCT within Monte Carlo (MC) simulations in order to fully understand their performance as part of CTA. 2) Examining the possibility of using the Cherenkov emission produced by local muons as a form of absolute calibration for GCT. 3) Investigating the use of clustering algorithms to efficiently discover sources in very high energy (VHE) data sets with a focus here on active galactic nuclei (AGN). This thesis therefore presents results on these topics.

In Chapter 2 an overview of cosmic rays and gamma-rays will be given. Initially focusing on the physical mechanisms involved in the production of high energy photons, such as inverse Compton scattering, Bremsstrahlung, synchrotron emission, pion production, nuclear de-excitation and particle annihilation. From here the important processes that lead to the attenuation of these photons, and more importantly their detection, will be covered. As the later work in this thesis will cover the search for AGN, a description of the current known population will be given along with a discussion of the main emission mechanism models.

Due to the high energy of gamma-rays, they can not be focused as with traditional telescopes. In addition, the atmosphere is opaque to the electromagnetic spectrum above UV wavelengths. Therefore in Chapter 3 the main observational techniques employed to observe these high energy events will be described. This will concentrate on the methods used by the *Fermi* space based observatory, which escapes the atmosphere in order to observe the gamma-ray sky at the cost of effective area, and the method used by Imaging Atmospheric Cherenkov Telescopes like CTA, which detect the secondary product of the interaction of high energy particles with the atmosphere. Finally a brief description of CTA will be given along with an overview of the GCT.

With the desire to reach a better understanding of the potential of GCT, Chapter 4 will cover the set-up of the telescope (SST-GATE) and camera (CHEC) within the MC framework. As there are currently two designs for CHEC that use either multi anode or silicon photomultiplier modules, a description of both will be given. From here, in order to evaluate the new GCT configuration, low level performance indicators such as charge resolution and trigger efficiency will be shown. Lastly, using

a dedicated production of data for a series of small arrays, high level performance indicators such as differential sensitivity and energy and angular resolution will be derived.

As with any instrument, a good understanding of the performance is required and must be constantly monitored to prevent inconsistent and poor quality results. Therefore, in Chapter 5, a study of the use of Cherenkov light from the well known spectrum of local muons as a method of absolute calibration for GCT will be presented. The method has been shown to be successful in other experiments such as H.E.S.S., VERITAS and MAGIC but it was not clear that it would be possible for the SSTs due to their reduced mirror area compared with that of previous instruments.

To gain a full understanding of the physical processes that are present in some of the most extreme environments in our universe, AGN and their associated relativistic jets, it is important to have a large and robust sample of objects. To this end, Chapter 6 and 7 will present ongoing work to use the clustering algorithm DBSCAN to efficiently detect sources within VHE data sets, using the highest energy events available (E_γ >100 GeV) within the *Fermi* large area telescope (LAT) instrument as a test bed. In Chapter 6 the initial study using Pass 7 Reprocessed *Fermi* data will be presented with Chapter 7 covering a further study using the improved Pass 8 data and a more advanced analysis algorithm.

Chapter 2

The Violent Universe

In this first Chapter, a brief introduction into the most energetic branch of astronomy will be given along with a description of the physical processes involved. A focus will be given to one particular source class that makes up the majority of known emitters at the highest energies, Active Galactic Nuclei (AGN), as these will become a point of interest in the last two Chapters of this work, emerging as the sole class of objects found using a clustering analysis of the gamma-ray sky.

2.1 Cosmic Rays and Gamma-Rays

In comparison to other fields of astronomy, gamma-rays represent a relatively new view of the Universe, one that is highly energetic and more akin to a vast particle accelerator than a sea of stars and dust. In 1912 it was confirmed by Victor Hess and Werner Kolhörster that Earth was being bombarded from above with highly energetic particles [75]. These particles, later described as cosmic rays by Millikan [92], sparked an interest as a possible new branch of astronomy. In 1927 they were identified as charged particles by observing that they were deflected by the Earth's magnetic field [52] and in 1934 deduced to be mainly positive through the measurements of a directional discrepancy between east and west. This deflection of cosmic rays along the Earth's magnetic field lines, in addition to the magnetic fields present in the Solar System and the Milky Way, unfortunately results in a loss of directional information, preventing a traditional view of the universe from being constructed[1]. However, work still continues to study the spectrum of cosmic rays incident at Earth [94], a full understanding of which could reveal fundamental information about the make-up of the universe[2].

In the production of the high energy particles that make up the cosmic ray flux, physical processes also lead to the emission of very high energy photons, known as gamma-rays. As neutral particles, they suffer no deflection from magnetic fields and can be traced back to their origin, therefore revealing the universe in the way that cosmic rays were once hoped to do.

Representing energies above 511 keV, the gamma-ray energy band covers a large

[1] This may not always be true, cosmic rays with enough energy can "resist" the deflection from the magnetic field. This is known as magnetic rigidity and allows for a possible correlation of some ultra-high energy cosmic rays (UHECR) with known astrophysical particle accelerators [57].

[2] Measurements of the cosmic ray flux reveal distinct features, or turning points, in the spectrum known as the "knee" and the "ankle". It is thought that different astrophysical objects are responsible for the spectrum between each feature with Supernova Remnants and other Galactic sources making up most of the emission below the knee at E$\sim 3 \times 10^1$5 eV, Active galactic nuclei and other extragalactic sources above the ankle at E$\sim 10^{18}$ eV and a combination of all sources between. See [80] for more detail.

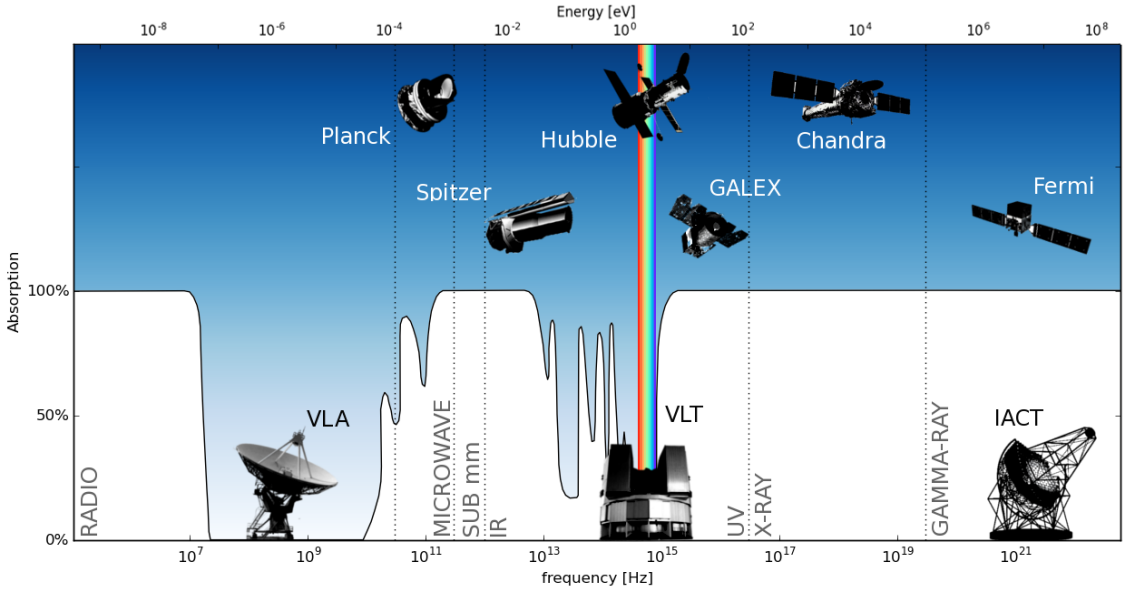


Figure 2.1: The electromagnetic spectrum from radio waves to gamma-rays. Also shown is the level of atmospheric absorption along with example observatories and satellites. While the atmosphere is technically opaque at gamma-rays, very high energy photons produce secondary effects within the atmosphere which allow for their detection by imaging atmospheric Cherenkov telescopes (IACTs).

proportion of the electromagnetic spectrum (see Figure 2.1), potentially extending to PeV energies [1]. Due to this high energy, gamma-rays cannot be focused in the same way as lower energy photons and instead we must rely on direct detections of their interactions with material. This can take the form of a heavy material such as tungsten foil in the case of the space-based gamma-ray observatory *Fermi's* large area detector (LAT) or in the case of ground-based gamma-ray detectors, the atmosphere itself. These instruments and techniques will be described in detail in the next Chapter, as it is first important to understand the physical processes that lead to the production and detection of gamma-rays.

2.2 Interaction and Production Processes

In general, gamma-rays are produced either through charged particles interacting with strong electric or magnetic fields (electromagnetic) or through hadronic processes. For reference a brief overview will be given of the most notable of these processes; for a more detailed description see [122], [77] or [108].

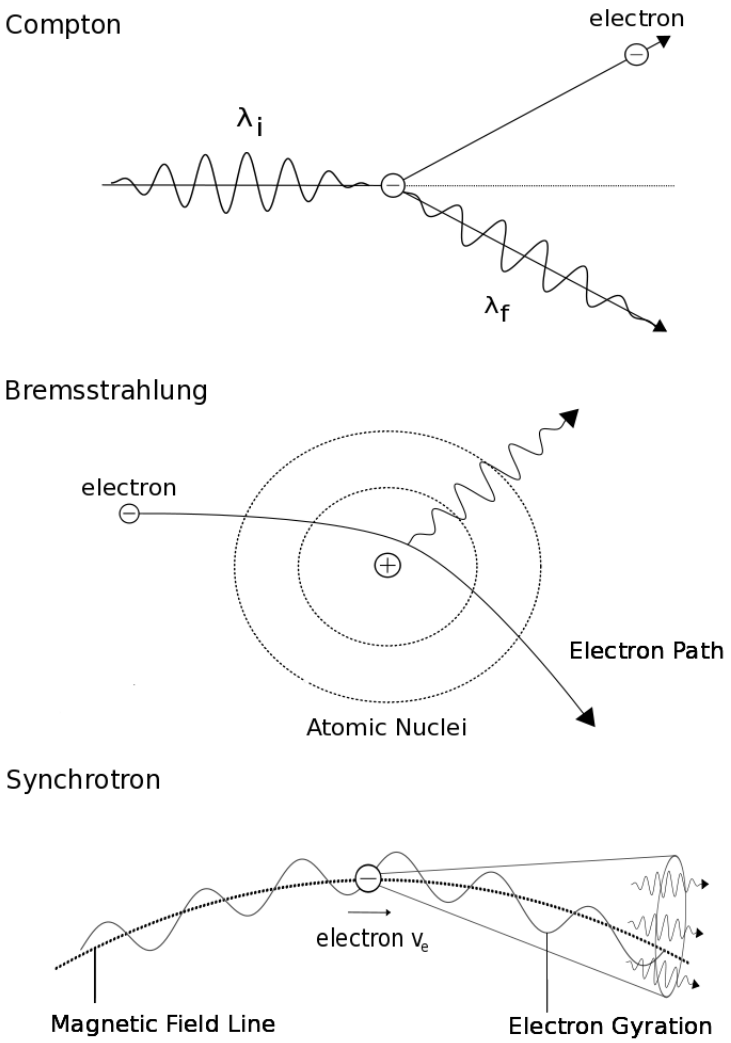


Figure 2.2: Example electromagnetic production processes of gamma-rays. Top: Compton scattering; while this example results in a lower energy photon being emitted (i.e. $\lambda_f < \lambda_i$), in the opposite case known as Inverse Compton Scattering, the electron is travelling at relativistic energies which results in an upscattering of the photon (i.e. $\lambda_f > \lambda_i$). Middle: Bremsstrahlung involves the production of a photon during the deflection of a relativistic electron in the Coulomb field of an atom. Bottom: Synchrotron radiation produced by a relativistic electron spiralling along a magnetic field line.

2.2.1 Electromagnetic

These processes are not necessarily unique to the gamma-ray energy regime, but instead are relevant when the energies of the participating particles or the field strengths are large. This situation is often found in the most extreme environments in the universe.

Inverse Compton Scattering

Compton scattering involves the interaction between an unbound electron and a photon. In its original definition, it represents a photon colliding with a stationary electron, imparting some energy and then being re-emitted as a photon with lower energy as can be seen in Figure 2.2 (top), where the change in energy can be represented as

$$\Delta E = \frac{E_i E_f}{m_e c^2}(1 - \cos\theta). \tag{2.2.1}$$

Here E_i is the energy of the incident photon, E_f is the energy of the scattered photon, m_e is the mass of the electron and θ is the angle of deflection of the photon. In the inverse case, a relativistic electron interacts with a photon, resulting in an "up-scattering" of the photon to a larger energy. In the rest frame of the electron, this equates to a resulting energy of

$$E_f \approx \gamma E_i(1 + \beta\cos\theta), \tag{2.2.2}$$

where $\beta = v/c$ is the velocity of the electron and γ is the Lorentz factor for the electron. Transferring to the observer frame this results in a relation of $E_f \sim \gamma^2 E_i$. This is often thought to be the dominant production method in gamma-ray astronomy, especially at high energies (See Section 2.4).

Bremsstrahlung

Electron bremsstrahlung, or "braking radiation" occurs when an electron is deflected by the strong electric field (Coulomb field) of an atomic nucleus; in this case the

amount of energy released is proportional to the acceleration caused by the deflection. This can be seen in Figure 2.2 (middle). In an astrophysical context this becomes an important mechanism in the interaction of relativistic electrons with atomic or molecular material. For example, in the production of diffuse Galactic gamma-ray emission though interaction of relativistic charged particles with the interstellar medium [125]. As will be shown later, this is also an important process in the development of electromagnetic air showers within the atmosphere.

Synchrotron Radiation

Charged particles, in this case electrons, are strongly affected by the presence of a magnetic field, resulting in a deflection along the field lines. If the electron has a component of motion perpendicular to the field direction, a correcting Lorentz force will cause a gyration about the field line. Therefore, the electron receives constant acceleration, with the energy loss and the particle momenta being conserved through the emission of a beamed cone of highly polarised synchrotron radiation, see Figure 2.2 (bottom). The resulting radiation is responsible for the majority of non-thermal emission, particularly from x-rays, but is generally observable as radio waves. As will be seen later, it is responsible for the low energy emission in active galactic nuclei. In the case where the magnetic field lines exhibit strong curvature, such as those found in pulsars, sufficient bending acceleration is present to produce "curvature radiation" in a process similar to bremsstrahlung.

2.2.2 Hadronic

Pion Production

When a relativistic hadron collides with matter, such as interstellar gas, the collision will often produce secondary particles. The most common example is the proton-proton interaction which results in the production of mesons, such as pions (and to a lesser extent kaons) in roughly equal abundance of positively charged (π^+), negatively charged (π^-) and neutral particles (π^0) [99]. These are the lightest mesons and have short lifetimes, decaying on time scales of 26 ns for the charged pions and

8.4×10^{-17} s for neutral pions, with the main decay products being photons, muons[3] and muon neutrinos:

$$\pi^0 \rightarrow \gamma\gamma, \qquad \pi^+ \rightarrow \mu^+ + \nu_\mu, \qquad \pi^- \rightarrow \mu^- + \bar{\nu}_\mu \tag{2.2.3}$$

It has long been theorised and recently confirmed at VHE energies that the decay of neutral pions into gamma-rays is responsible for the majority of the diffuse VHE emission along the Galactic plane [2].

De-Excitation of Atomic Nuclei

Through processes such as photo-ionisation collisions, atomic nuclei can be moved to a higher energy state. The resulting decay of this energetic state will result in the production of a photon, which as nuclear states have energy spacings of ∼MeV, will be observable in the lower gamma-ray energy regime. One area where this takes place is near the central engine of an AGN in the gravitationally bound gas clouds, therefore producing distinct atomic lines in the areas known as the broad line and narrow line regions. The presence of these lines is extremely useful for determining the distance to the AGN through the measurable redshift of the known atomic lines.

Annihilation

The annihilation of particles with their corresponding anti-particles is also responsible for the emission of gamma-rays. Take, for example, the lightest possible particle-antiparticle pair, the electron and positron. Knowing that the mass of an electron is ∼511 keV/c^2, then the total energy produced will be 1.022 MeV. In order to conserve momentum in the system, two photons must be produced each with 511 keV in the rest frame of the annihilation which results in a measurable line at ∼511 keV [103].

In addition to this, a large effort is also being applied to the search for possible spectral lines resulting from the annihilation of dark matter (DM) particles. Many

[3]Muons can undergo further decay into either an electron or a positron and an electron neutrino. See Chapter 5 for more on muons.

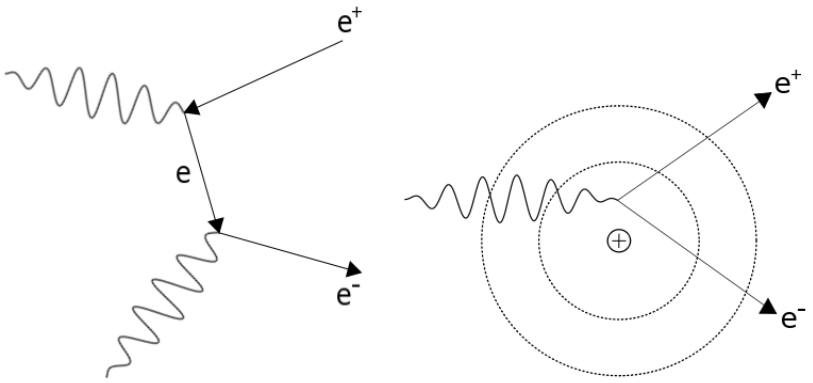


Figure 2.3: Left: example Feynman diagram of two photons interacting to produce an electron positron pair, Right: a photon interacting within the Coulomb field of an atom resulting in the production of an electron positron pair.

models predict that the mass of DM particles is large enough that this should result in a line in the gamma-ray regime [6].

2.3 Detection and Attenuation Processes

It is important to now consider the physical processes that govern the interactions of gamma-rays that lead to their attenuation or eventual detection.

2.3.1 Interactions With Light

One of the largest attenuation processes for VHE gamma-rays is the interaction with other light, especially abundant low energy photons. When two photons with a total energy greater than the rest mass of two electrons, such that

$$E_{\gamma+\gamma} > 2m_e c^2 \tag{2.3.4}$$

where m_e is the rest mass of an electron and c is the speed of light, the photons will undergo electron-positron pair production, effectively eliminating the gamma-ray photon,

$$\gamma + \gamma \rightarrow e^- + e^+ \tag{2.3.5}$$

An example of this can be seen in the left panel of Figure 2.3. While this happens largely in areas with a high density of low energy photons, particularly at the production sites of gamma-rays, it is also responsible for one of the most important attenuation processes of distant sources. In the space between galaxies, there exists a steady background of photons known as the Extragalactic Background Light (EBL) of which a component is thought to have been produced by the first stars to have formed in the early universe [15]. This poses an important cosmological question regarding the amount of EBL, which can be measured by considering the optical depth of VHE photons as a function of redshift, and therefore by its effects on the high energy spectra of distant sources, e.g. [90].

2.3.2 Interaction with Matter

In a similar process to the interaction with light, a gamma-ray photon can interact within the Coulomb field of matter to again produce an electron-positron pair, such that

$$\gamma + \text{nucleus} \rightarrow \text{nucleus} + e^- + e^+. \tag{2.3.6}$$

An example of this can be seen in the right panel of Figure 2.3. While this can lead to an attenuation of the gamma-ray signal, the amount of matter in the path between the source and Earth is small and leads to a minimal effect. In comparison, the Earth's atmosphere presents a very dense medium and results in most gamma-rays interacting in this way to produce electron positron pairs. It is this process that, instead of preventing their detection, allows for the observation of VHE photons by ground-based instruments. Due to the energetic nature of the resulting electron and positron, a particle shower is generated with an associated flash of Cherenkov radiation. This will be described in detail in Chapter 3. Before this, let us consider an example astrophysical source in order to further understand the importance of these processes.

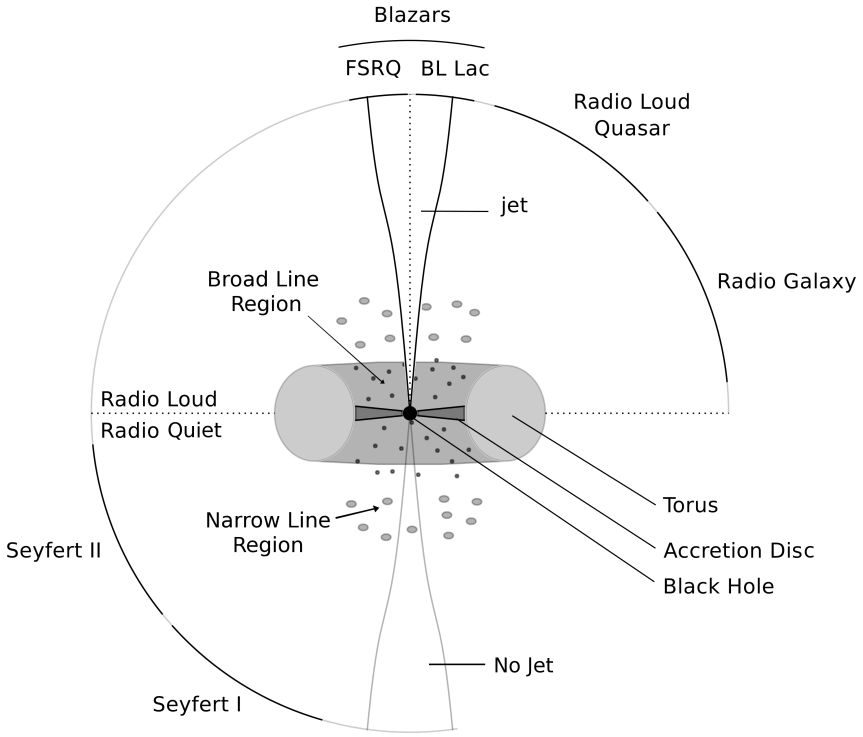


Figure 2.4: The orientation-based unification theory of AGN in which classification is based firstly on the presence of a relativistic jet and secondly on its orientation with respect to the observer. Adapted from [117]

2.4 The Extragalactic Sky and Active Galactic Nuclei

Active Galactic Nuclei (AGN) are galaxies that contain a compact, bright and often variable core. The engine at the centre of each AGN has for a long time been known to be a super massive black hole (SMBH), fed by a hot accretion disc and surrounded by a dusty torus. As matter collapses inwards towards the SMBH, a hot plasma is formed. However, not all this material is destined to join the singularity and is instead directed towards the poles and ejected with immense energies. The result is the formation of spectacular jets containing relativistic particles that extend millions of light years perpendicular to the plane of the galaxy. The resulting luminosity is comparable to or dominant over the emission from the host galaxy. A general unified theory of all AGN has been based on their luminosity and the geometry of the AGN with respect to the line of site to Earth [22], although this is often considered

an oversimplification [82]. In Figure 2.4, a side-on view of an AGN can be seen, showing the central SMBH, the accretion disc, the dusty torus and the jet. Here radio quiet galaxies, i.e. Seyfert I and Seyfert II, generally have lower luminosities and often exhibit spectral lines (Seyfert II galaxies only exhibit spectral lines from the narrow line region, with the assumption that the broad line region, situated closer to the core, is obscured). Radio loud galaxies, consisting of $\sim$10% of the total galaxy population, are defined by the presence of a relativistic jet which acts as a production site for high energy photons. Radio galaxies represent a side view of the galaxy where the core is obscured by the dusty torus but jets and lobes are visible. At acute angles the inner core region becomes visible, exposing the broad line region and often greater variability. Finally blazars, described as looking "down the barrel of the gun", with the jet orientated towards Earth, result in large variable luminosities. Blazars are divided into two sub-classes: BL Lacs, which are named after the first object of this type (BL Lacertae) and have a completely featureless spectrum, often making distance measurements difficult and Flat Spectrum Radio Quasars (FSRQs) which differ from BL Lacs in the presence of observable line emission.

In this situation, different orientations result in alternative classification, a concept well expressed in [122]

> "In the unified picture of AGN it is probable that all AGN, at some level, are emitters of high energy gamma-rays and that only some are detectable is an accident of orientation, of geometry, not physics" - Weekes

And it is indeed true that here the specific interest is in AGN that are observable in gamma-rays. In the above classification, it is obvious that the least numerous AGN types, based solely on their geometry, are blazars in which the jet is aligned within a few degrees of the observed direction. However at high energies, this ratio is completely shifted. At TeV energies, the detected galaxies are almost exclusively blazars.

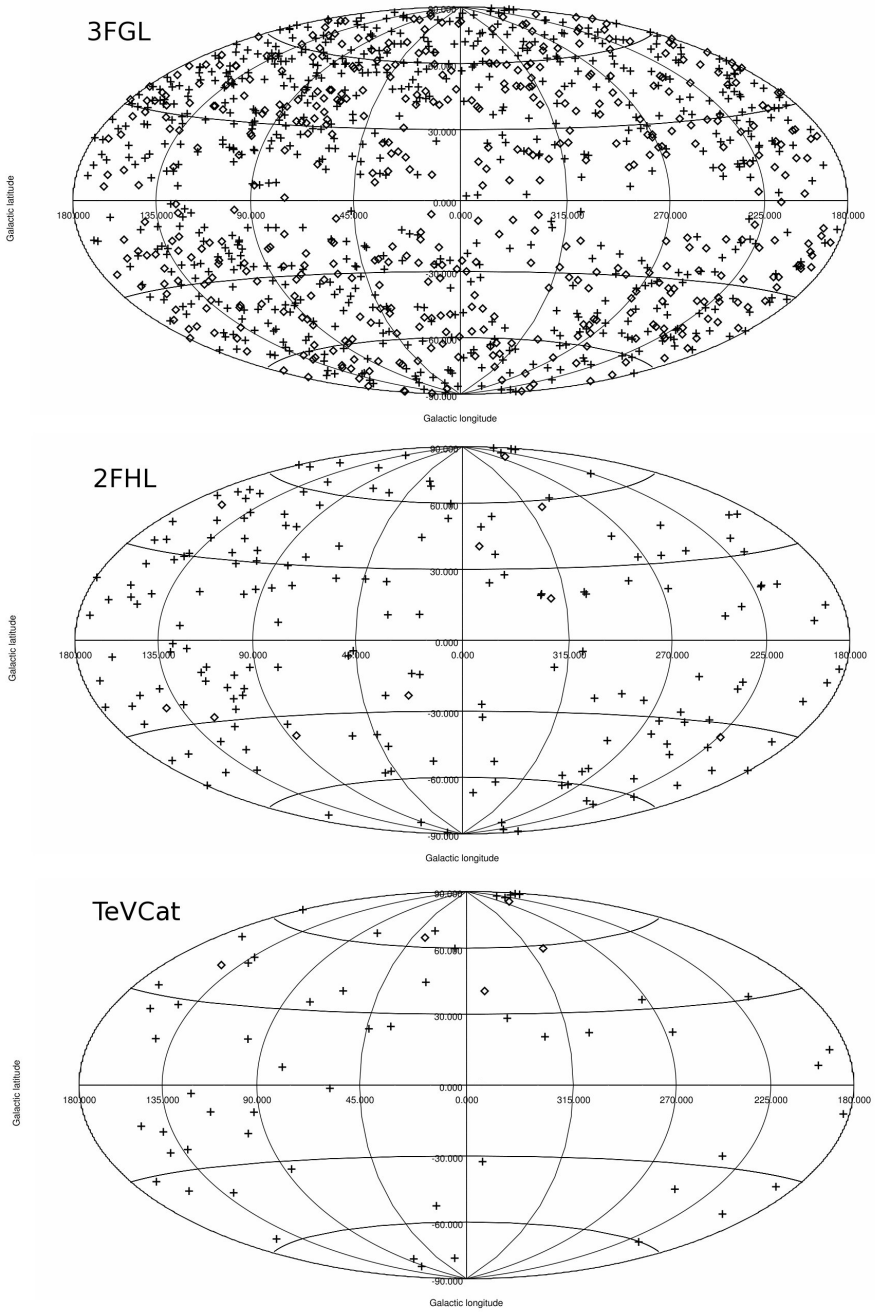


Figure 2.5: Known AGN within the 3 year Fermi LAT point source catalogue (10 MeV - 300 GeV, top) AGN within the 2nd Fermi catalogue of hard sources (50 GeV - 2 TeV, middle) and known VHE AGN within the TeVCat catalogue as of the 26th May 2016. Here black diamonds represent FSRQs and black crosses represent BL Lacs.

2.4.1 VHE Blazars

Due to their great luminosities, together with their emission mechanisms and the effects of relativistic Doppler boosting (as will be shown shortly), blazars constitute the most numerous source class observed in the extragalactic sky at GeV energies and above. The space-based gamma-ray observatory Fermi has detected 660 BL Lacs and 484 FSRQs (along with 573 unclassified AGN) within its 4-year point source catalogue (3FGL [10]) between 100 MeV and 300 GeV. More recently, the Fermi collaboration also published the 2nd catalogue of hard Fermi-LAT sources (2FHL [11]) which contained a reduced number of blazars, 180 BL lacs and 10 FSRQs in their upper-most observable energy range of 50 GeV - 2 TeV. Finally, moving to ground-based VHE instruments, the number of BL Lacs drops to 56 and FSRQs to 5 based on observations from experiments such as H.E.S.S. [78], VERITAS [121], MAGIC [86], Durham MK6 [23], CANGAROO [93] and Whipple [83]; these sources are all recorded in the TeVCat catalogue[4]. The positions of all these AGN can be seen in Figure 2.5. One reason for the dramatic drop between 2FHL and TeVCat is due to the observation requirements: While *Fermi* is able to scan the entire sky every 3 hours, ground-based instruments are restricted to pointed observations, operating under dark sky conditions and are weather dependent, reducing the possible observation time to a duty cycle of around 10%. In addition, ground-based instruments perform a lot of follow-up observations based on multi-wavelength alerts and rely on active states (i.e. larger luminosities) to detect some of the fainter objects, leading to a somewhat biased sample. To understand this further, it is important to consider how it is possible to observe such extreme luminosities, along with the emission mechanisms that lead to gamma-ray production.

2.4.2 Relativistic Jets

There are two main pieces of evidence that lead to the assumption that the gamma-ray emission from AGN originates within the jet. Firstly, at lower energies, between

[4]the TeVCat catalogue can be found at http://tevcat.uchicago.edu/ Date of access: 26th May 2016

10 and 20% of the emission that is observed from these sources is polarised [35]. Recall from Section 2.2.1, that emission from synchrotron radiation is known to be polarised, implying the presence of a large population of relativistic particles [34]. Secondly, the gamma-ray emission appears to originate, based on the observed rapid time variability, from a compact region[5] [68]. Given the observed luminosities of some of these objects, in order to prevent a situation in which the VHE emission Compton scatters or pair produces, it is necessary to assume relativistic beaming along a jet. In this assumption, a bulk relativistic motion is introduced with an associated bulk Lorentz factor (Γ) and associated Doppler factor (δ):

$$\Gamma = \frac{1}{\sqrt{1-\beta^2}}\ , \qquad \delta = \frac{1}{\Gamma(1-\beta cos\theta)}, \tag{2.4.7}$$

where $\beta = v/c$ is the velocity and θ is the angle between the direction of motion and the line of site of the observer. In some recent observations, Doppler factors of 60-120 have been calculated (e.g source PKS 2155-304) [17]. This Doppler factor helps explain the following problems:

1. **Small Emission Volume:** If the emission region is defined as $D \sim dt_{obs}c$ where dt_{obs} is the observed time variation, the gamma-ray photon density is so large that they would undergo pair production on the soft photon component and the signal would be greatly reduced. By introducing the Doppler factor a shift is applied of the form $dt_{source} = \delta \cdot dt_{obs}$, allowing for a larger emission region.

2. **High Frequency:** The observed high-frequency peak requires large maximum energies to be achieved within the jet. By including the Doppler scaling, the frequency is reduced in a similar way to the time scale, $\nu_{obs} = \delta \cdot \nu_{source}$

3. **Extreme Luminosity:** Finally, the required source luminosity is drastically reduced, with the observed value boosted by a factor of δ^4. This helps reduce

[5]Such short time scale variation would be physically impossible over large areas due to the time needed to traverse the emission region.

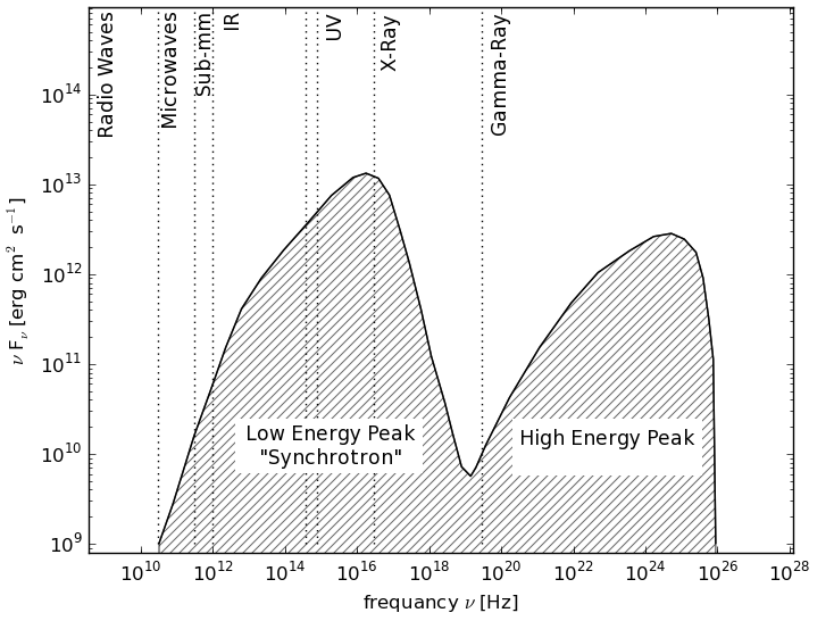


Figure 2.6: An example spectral energy distribution for a blazar, which exhibits two distinct peaks. The first peak, which spans from radio waves and into x-rays, is generally agreed to originate from synchrotron emission from relativistic electrons within the jet. The second, high-energy, peak highlights the importance of observations in the gamma-ray spectral range as the cause of this peak is still a subject of debate. Most theories describe the process as either hadronic or leptonic in nature (or often a combination of both).

the optical depth for pair production, resulting in an optically thin emission region and therefore allowing generated gamma-rays to escape.

Having arrived at an understanding of how such a luminosity is possible, the current models for emission within the jet are considered. The power output for blazars is expressed in a spectral energy distribution (SED) that shows the measured power at each frequency. An example for a typical blazar can be seen in Figure 2.6 and shows two distinct peaks[6]. The low frequency peak is generally agreed to originate from incoherent synchrotron emission by relativistic electrons within the jet, spanning frequencies from radio to X-rays (the position of the peak is determined by the efficiency and the cooling process). The origin of the second peak, which calls

[6]Some blazars also have contributions from other emission regions, e.g. from the accretion disc, which affects the spectral shape

for observations with gamma-rays is still contested, but is widely accepted to arise from either lepton (Section 2.2.1) or hadron (Section 2.2.2) based emission models.

2.4.3 Leptonic I: Synchrotron Self Compton (SSC)

In an environment containing the relativistic electrons and the resulting synchrotron photons that make up the first SED peak, it is probable that there will be a certain amount of inverse Compton scattering of the synchrotron photons by the same electron population that produced them (termed synchrotron self Compton). In this situation the photons will be boosted up to energies close to that of the participating electron. One of the most simplistic theories for SSC is the one zone model which describes a population of relativistic electrons injected into a spherical emission region which moves with a relativistic speed along the jet [37]. This model is often favoured as it incorporates the least degrees of freedom, however it suffers from an inability to explain so called "orphan flares", i.e. where a flare is only visible in one of the SED peaks or there is a delay between the low and the high energy component. In order to solve this, multi-zone emission models are often applied, in which the different energy peaks are produced in different regions [38].

2.4.4 Leptonic II: External Compton Radiation (ECR)

Instead of boosting the synchrotron photons produced within the jet, this model relies on a large density of soft photons produced external to the jet, for example from the accretion disc, torus or broad line region, that can be boosted to the energies seen within the second peak. By assuming a population of external photons not produced within the bulk flow, the result is a constant field of seed photons which are then up-scattered and an increase in the rate by which the relativistic electrons lose energy [110].

2.4.5 Hadronic

In these models, the source of the photons in the high-energy peak comes instead from the accelerated protons in the jet. These protons, with energies of up to

10^{18} eV, interact with the soft photons and undergo photo production, producing pions which then decay into gamma-ray photons, or alternatively emit photons via synchrotron radiation. Hadronic models are often motivated by the observed proton flux in the cosmic ray spectrum, providing AGNs as the source of protons up to 10^{20} eV. A potential confirmation of the hadronic emission model could come from the detection of neutrinos which are known to be produced in proton interactions (recall Section 2.2.2) [13].

2.4.6 The Path to Understanding

At present observations are unable to determine the exact emission mechanism, with different AGN requiring different emission models. The complication arises from deconvolving the intrinsic emitted VHE energy spectra and the effects of complex absorption and cascade processes in the host galaxy along with those within extragalactic space (e.g. EBL). The current sample of VHE AGN is generally considered to be incomplete and to suffer from a strong observation bias towards active states and increased luminosities from large Doppler boosting. This is mainly a result of the sensitivity limits of current instruments and the observation strategy adopted, relying heavily on follow-up of multi-wavelength alerts. It is however foreseen that the planned new ground-based gamma-ray observatory, the Cherenkov Telescope Array (CTA), with an order of magnitude improvement in sensitivity, will help to rectify this. Providing deeper observations of selected AGN and aiming to expand the sample of known blazars, especially increasing the range of redshift covered, will allow a reliable luminosity function to be constructed. In order to predict the success of CTA in this respect, an initial prediction of the number of detectable AGN has been made in [111]. Here, a sample of lower frequency extragalactic sources with known redshifts has been taken from the 2 year point source Fermi catalogue [9]. Using the published spectral parameters, a conservative estimate is that 170 AGN should be detected, allocating a maximum of 50 hours of observation per source[7].

[7] A less conservative approach can be taken by removing the selection criterion that the redshift must be known, leading to a predicted 370 detected sources.

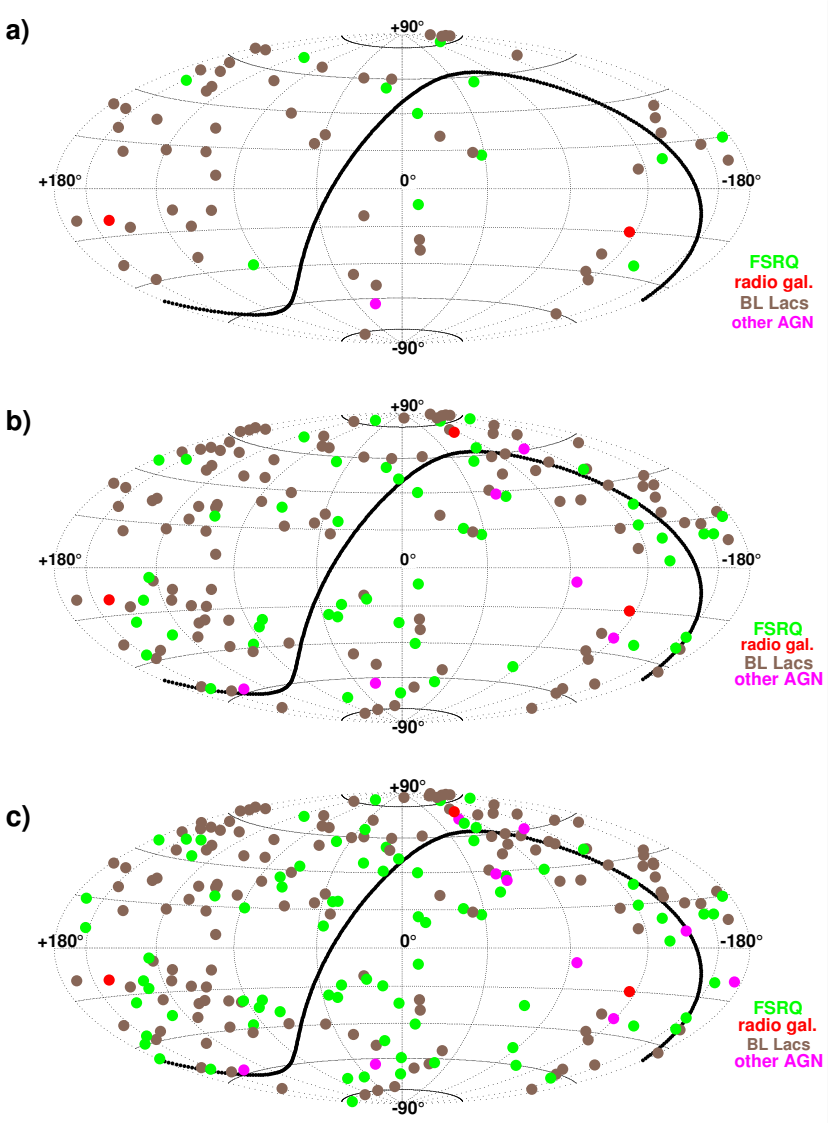


Figure 2.7: Predicted number of AGN detected with 5 hours (a), 50 hours (b), and 150 hours (c), of observation with CTA, resulting in 69, 170 and 230 sources respectively. Based on the Fermi 2 year point source catalogue (2FGL [9]), figure taken from [111].

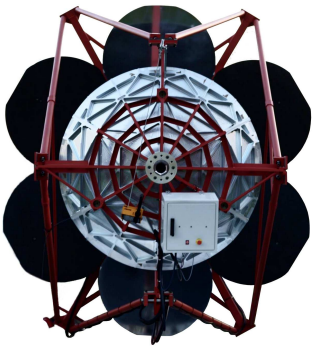
Increasing the exposure time to 150 hours results in a prediction of 230 sources detectable in less than 10 years. The predicted sky maps produced based on these assumptions can be seen in Figure 2.7.

While a method utilising observations at lower frequencies can result in a large number of newly-detected AGN at TeV energies, it still introduces further observational bias. The most reliable method to avoid this is to perform a completely blind survey of the extragalactic sky, which, as well as discovering as yet unknown AGN, may detect additional unexpected phenomena. This is in fact one of the goals of CTA, which can reach sensitivities down to 20 mCrab at 5σ in a 30 minute exposure, which is around the level of the faintest currently detected VHE AGN. Assuming a field of view of 5°, this equates to an ability to cover a quarter of the extragalactic sky in a quarter of the available observation time in a year [61]. Unfortunately, due to our limited knowledge of the properties of AGN, especially at VHE, it is difficult to produce a robust estimate of the expected return of such a survey.

In this Chapter, an overview of some of the most important physical aspects of gamma-ray astronomy has been presented, along with an example of how observations can reveal important information about acceleration processes in AGN. Throughout, the space based observatory *Fermi* and the future planned ground based observatory CTA have been mentioned. In the following chapter, these will be described in greater detail, especially how they are able to detect photons of such great energy.

Chapter 3

Observational Techniques and the Future of γ-ray Astronomy



In this Chapter, a more detailed description of the observational methods used in gamma-ray astronomy will be given, focusing on two main observatories. First the space-based observatory *Fermi*, which uses a silicon tracker to directly detect pair production caused by gamma-rays in its large area detector (LAT), and secondly, the planned ground-based observatory, the Cherenkov Telescope Array (CTA) which will draw on the experience of current ground-based experiments to reach sensitivities an order of magnitude greater than currently possible.

3.1 Introduction

In the previous Chapter the primary mechanism by which gamma-rays interact with matter was introduced, i.e. pair-production. In this Chapter, these interactions will be covered in more detail, specifically concerning the methods of detection of VHE gamma-rays. The two projects that will be focused on are the space-based observatory *Fermi* and the future ground-based observatory, the Cherenkov Telescope Array (CTA).

In Section 3.2 an overview of the *Fermi*-LAT instrument will be given and in Section 3.3 the physics behind extensive air showers and their resulting Cherenkov radiation along with the imaging techniques involved will be presented. In Section 3.4 the Cherenkov Telescope Array will be introduced, moving then to a description of the Gamma Cherenkov Telescope in 3.5.

3.2 Space Based: Fermi

The space-based satellite named after Enrico Fermi (previously GLAST) has been observing the whole gamma-ray sky between 100 MeV and 300 GeV since its launch in 2008[1]. By operating above the atmosphere, it is able to observe directly the pair production of gamma-rays. However, the drawback is that due to the cost and mass restrictions of launching a satellite into orbit, the maximum effective area is only around $\sim$0.8 m^2 at 100 GeV and therefore its sensitivity is limited. The satellite has two main instruments, the gamma-ray burst monitor (GBM) which observes transient sources [91] and the large area telescope (LAT) [27], the latter being the more sensitive and the instrument that will be focused on here.

The LAT instrument consists of a 4×4 array of precision converter-tracker modules, each in turn with 16 repeating planes which contain a high Z material, in this case tungsten, in which the gamma-rays have a high chance of converting into electron positron pairs. These are followed by two layers of silicon strips which read

[1]More recently extended to around 10 MeV $< E_\gamma <$3 TeV with a new reprocessed data set release.

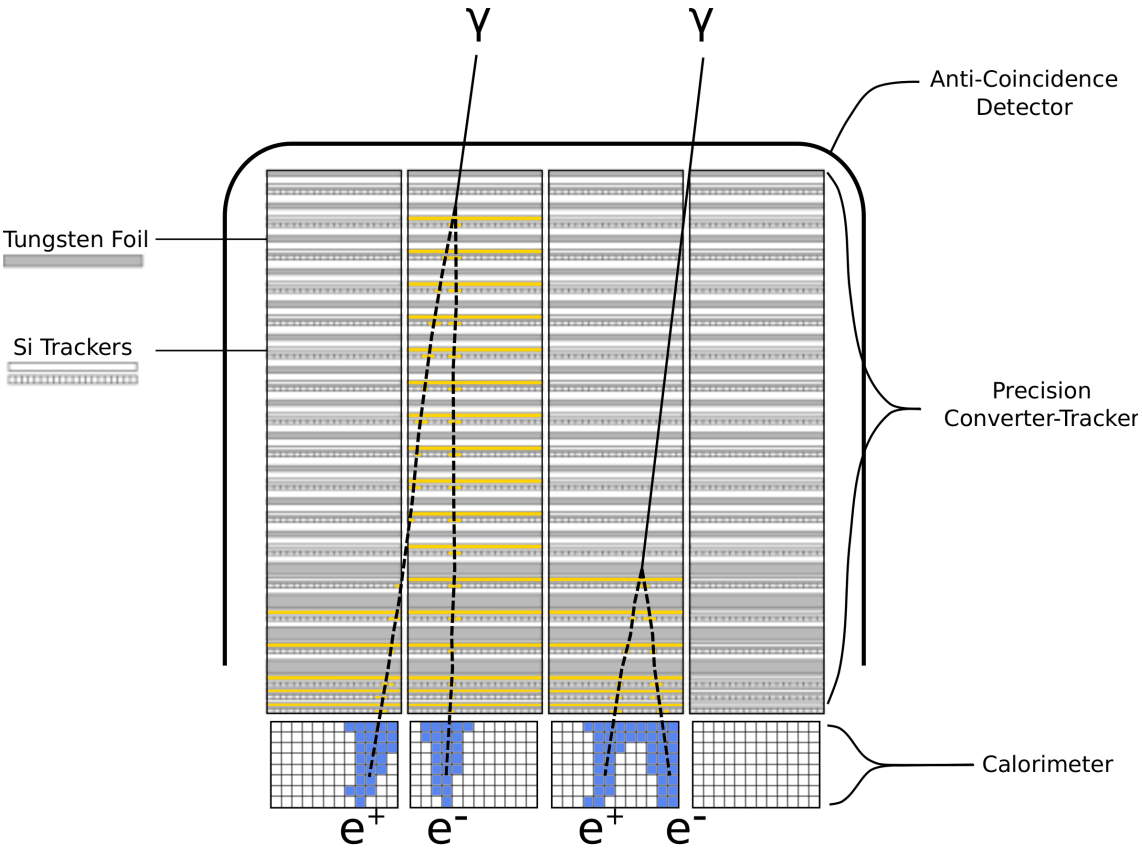


Figure 3.1: Simplified diagram of the *Fermi*-LAT instrument, showing the outer anti-coincidence layer, the tungsten foil conversion layers, the silicon detectors which have two layers for reading out X and Y coordinates and finally the calorimeter which measures the energy of the electron positron pair. As a gamma-ray photon enters the detector it undergoes pair production within one of the tungsten foils (it may interact in other areas but with lower probability). The path that the electron positron pair takes is recorded by the silicon layers, reading out the X and Y coordinates at each layer. Finally the energy is recorded in the calorimeter and the event is reconstructed using this and the directional data recorded by the silicon strips.

out in the x and y planes to allow for positional reconstruction of passing positrons or electrons by the production of ions. At the back of the instrument the energy of the e^-e^+ pair is recorded, along with the lateral shower development, via the 96 CsI(TI) crystal scintillators arranged in 8 layers which make up the calorimeter. Surrounding the instrument is a segmented plastic scintillator anti-coincidence detector that is designed to have a 99.97% detection efficiency to charged particles [27]. A simplified version of the LAT instrument can be seen in Figure 3.1.

As a photon enters the LAT detector the chances are high that it will undergo pair production, with an enhanced probability that this will occur within one of the tungsten foil layers. As the e^-e^+ pair passes through the rest of the instrument, it is subject to multiple scattering and deflection resulting in bremsstrahlung production, which all adds up to an uncertainty in the reconstructed direction of the incident gamma ray. To quantify this, the probability distribution for the reconstructed direction of the incident gamma-ray, as expected from a point source, is calculated. From this the point spread function (PSF) is defined as the containment radius that covers a certain percentage of events (usually defined as either 68% or 95%). Therefore, where the gamma-ray converts and how well it is detected by the silicon layers greatly affects this directional uncertainty. In addition to this, the tungsten conversion layers are divided into two categories, "FRONT" and "BACK", based on different desired optimisations. The first 12 tracking planes have thin layers of foil that optimise the PSF for low energy events (reducing multiple scattering) while the last 4 are around 6 times thicker with the aim of increasing the effective area for higher energy events (due to the increased amount of converting material the events are more likely to pair produce). Therefore the events recorded by *Fermi* LAT are divided into two event classes, FRONT and BACK, depending on the location of the conversion of the gamma-ray within the tracker with the former having an inherently better PSF [7].

The trigger for the LAT consists of a coincidence measurement from the silicon trackers, the calorimeter and the anti-coincidence layer (which acts to veto events that arise from cosmic rays). A large amount of event reconstruction is performed on board *Fermi* to reduce the data rate that needs to be transmitted down to Earth

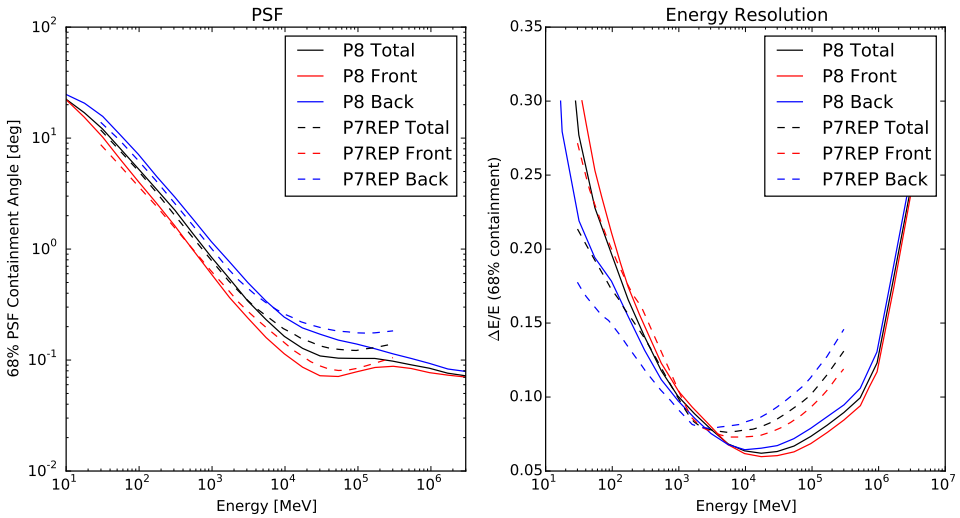


Figure 3.2: A comparison of the Pass 7 reprocessed and Pass 8 instrument response functions (IRFs), in this case the point spread function (left) and the energy resolution (right). Pass 8 represents a general improvement over the previous IRFs and has also increased the energy range of the available data.

to ∼1 Mbps. Further analysis and event quality definition is performed later by the *Fermi* LAT collaboration. There has been a series of releases containing processed and reprocessed event data along with updated instrument response functions, with the two most recent being the Pass 7 reprocessed [41] and Pass 8 [28]. These are needed in part to correct for degradation in the instrument over its long operation, for example the expected 1% per year signal loss from the calorimeter crystals due to irradiation, and to implement improvements in event reconstruction. The updates made during Pass 8 were the most dramatic, with the implementation of a completely new event reconstruction software which led to an improved PSF, energy resolution and also to an extension of the upper energy range of *Fermi* LAT to ∼3 TeV. See Figure 3.2 for a comparison of the PSF and effective area for the two different event processing schemes.

The specific instrument response functions for *Fermi* not only differ in the use of FRONT or BACK converting events, but also in respect of the data quality for specific analyses as determined by the *Fermi* LAT collaboration. These include SOURCE class events which are to be used with all basic analysis, CLEAN, which

have a lower background at higher energies and are more suited to hard spectrum sources at high galactic latitudes, ULTRACLEAN events which contain the lowest level of possible contamination and are used to check for cosmic ray systematics and TRANSIENT events, which contain less restrictive cuts to include more data [7].

3.3 Ground Based: Imaging Atmospheric Cherenkov Telescopes

The main disadvantage of *Fermi* is its small effective area that reduces its sensitivity to high energy photons. In order to improve this, one could imagine launching a satellite with a larger detector. Unfortunately, unless a lighter technology can be utilised, the cost becomes too great. However, the pair production of gamma-rays that occurs within the LAT detector also happens within the Earth's atmosphere in a process that leads to a large shower of energetic particles, known as an Extensive Air Shower (EAS). These showers have different components depending on the type of initiating primary particle. In the case of a gamma-ray or an electron, the shower development is governed by electromagnetic effects and is therefore defined as an electromagnetic air shower. If instead the primary is hadronic in nature (protons or heavier nuclei), then the shower development will consist of both electromagnetic and hadronic components. In the following sections we will see how these showers, along with the imaging of their by-products, can lead to the use of the atmosphere as part of a large detector.

3.3.1 Electromagnetic Air Showers

When a high energy photon enters the atmosphere an interaction will occur in the presence of the Coulomb field of an atmospheric atomic nucleus leading to the production of an electron and positron. In the next stage, the $e^{-}e^{+}$ pair will undergo deflection by further atomic nuclei resulting in an energy loss in the form of bremsstrahlung radiation, and will continue to do so until the energy has been reduced to a critical energy. At the same time the bremsstrahlung photons will also

pair produce resulting in more and more secondary particles. This can be easily displayed using the basic model proposed by Bethe and Heitler [89] in which several assumptions are made. Firstly, only the interaction processes of bremsstrahlung and pair production are considered, for which the radiation and conversion lengths are set to be equal (X_0). Finally, the energy is assumed to be divided equally between particles at each interaction. Therefore the whole process can be visualised as in Figure 3.3.

In the above simplified model, the number of shower particles at each level doubles (including bremsstrahlung photons, electrons and positrons) and therefore after n radiation lengths the total number of particles will be 2^n. Each of these particles, assuming that the energy is divided equally among them, will have an energy of $E(n) = E_0/2^n$, where E_0 is the energy of the primary. At some point within the shower, the energy of each particle will drop below a critical energy E_c at which point the production of new particles will sharply drop off. For example the minimum energy for pair production will occur at 1.02 MeV, i.e. twice the rest mass of an electron. Therefore a shower maximum can be defined with $N_{max} = 2^{n_{max}} = E_0/E_c$ and a depth of

$$t_{max} = n_{max} \cdot X_0 = \frac{\ln E_0/E_c}{\ln 2} \cdot X_0. \tag{3.3.1}$$

This is of course a simplistic representation of the true system; to arrive at a more accurate prediction coupled cascade equations should be used, a good description of which can be found in [31].

3.3.2 Hadronic Air Showers

In the case of a proton (or other heavier nuclei), the governing process is the inelastic scattering with atmospheric nuclei resulting in the production of mesons (and protons), of which the predominate population is the lightest known meson, the pion. Recall from Section 2.2.2 the short lifetimes and decay products of each of the pions (π^0, π^+, π^-). From this point the shower continues with a hadronic core continuing the above process. The photons produced from the decay of π^0 mesons result in secondary electromagnetic air showers, while a large proportion of the long-

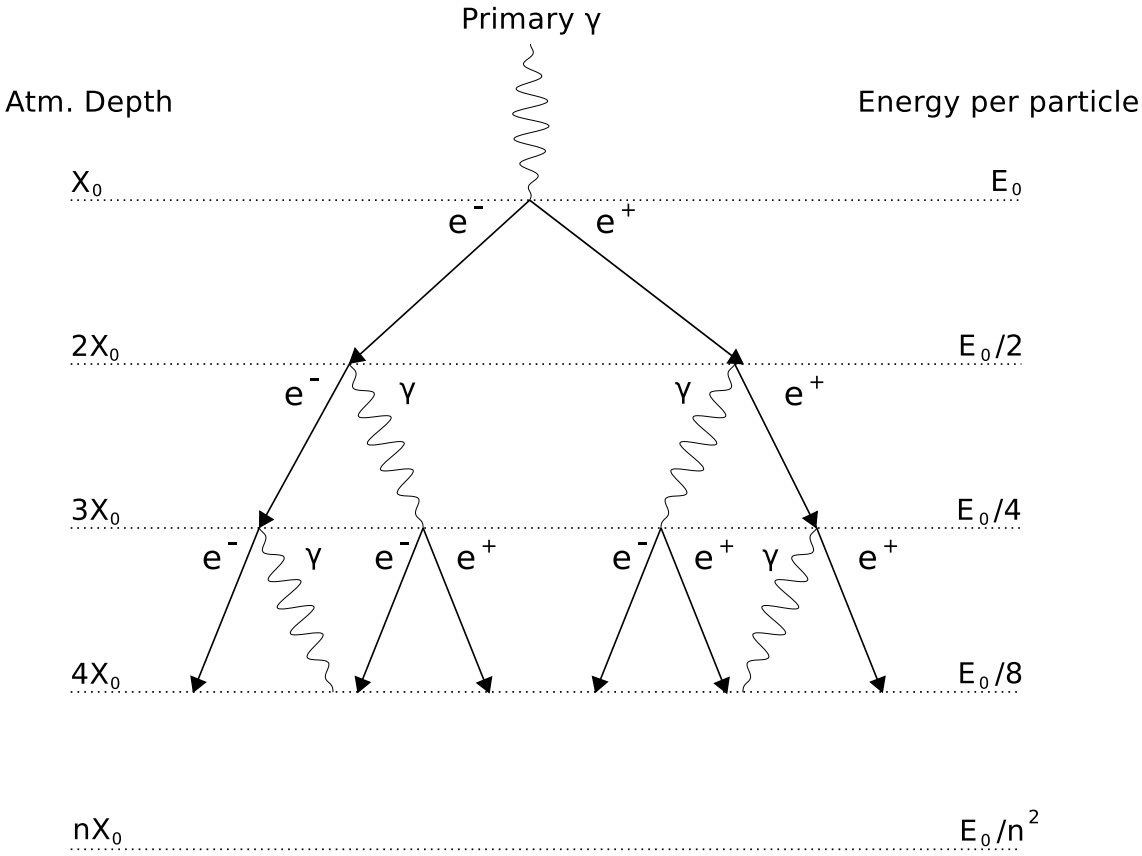


Figure 3.3: Extensive air shower generated by a primary gamma-ray which begins as production of an electron positron pair. Based on the Heitler model in which the conversion and radiation lengths are equal and the energy is divided equally at each interaction. Adapted from [89]

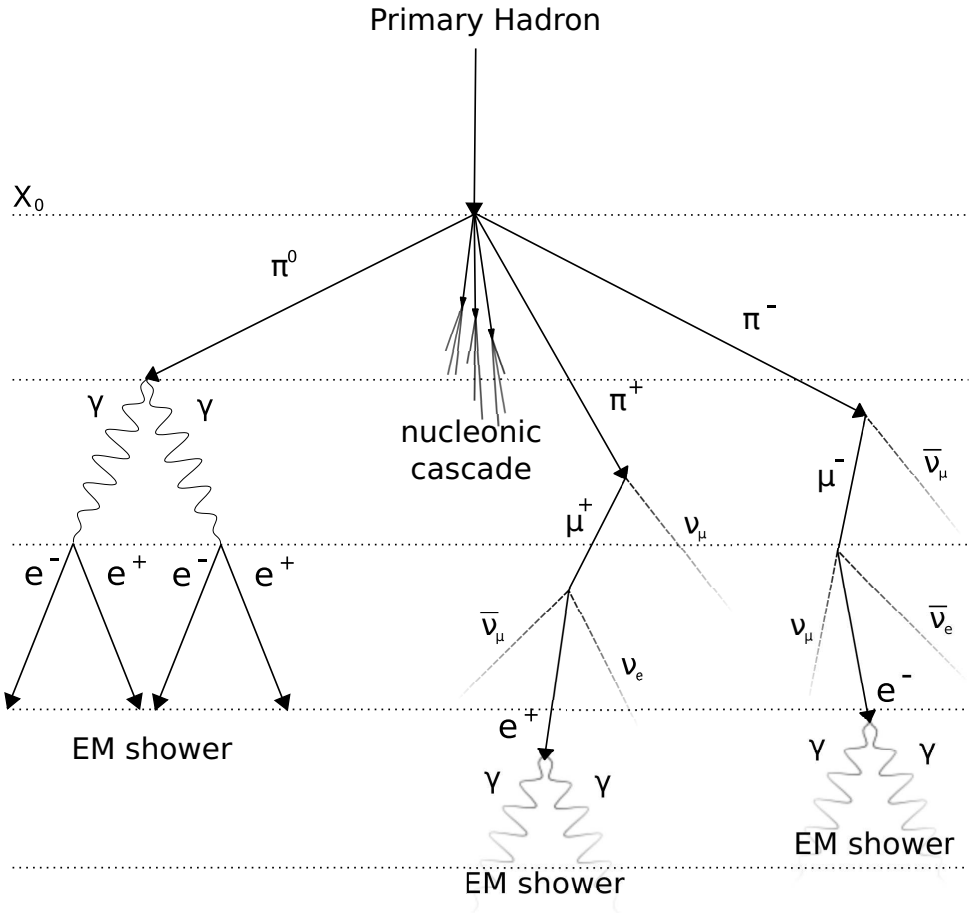


Figure 3.4: Extensive air shower generated by a primary hadron which produces a hadronic cascade along with pions which decay rapidly into photons, muons and neutrinos resulting in secondary electromagnetic showers. Figure adapted from [105]

lived muons and muon neutrinos produced by the decay of charged pions escape the shower region. The loss of energy from the shower from the muons and neutrinos, along with that dissipated via the inelastic collisions, removes ∼1/3 of the total energy, therefore requiring hadronic shower primaries to be of much larger energy than the corresponding electromagnetic shower primaries in order to produce a similar level of observable Cherenkov radiation (See Section 3.3.3). The process of inelastic scattering results in transverse momentum of the order of ∼100 MeV which in turn results in a greater lateral extension of the resulting shower in comparison to those induced through purely electromagnetic processes. For a simplified representation of a hadronic shower see Figure 3.4.

3.3.3 Cherenkov Radiation

As charged particles, such as the electron and positrons discussed above, move through matter, their electric field acts to polarise the medium through which they are propagating. After the charged particle has passed, the polarisation states collapse. In the case of a slow moving particle, this has no effect as the polarisation is symmetrical, i.e. there is no overall electromagnetic field. However, if the particle is moving through the medium at speeds faster than the local speed of light, then the polarisation is no longer symmetric in the direction of motion. In this case the collapse of the polarisation can only be achieved with a pulse of electromagnetic radiation known as Cherenkov radiation. These pulses then propagate outwards and a cone of constructive interference, governed by Huygens' principle, is generated causing a Cherenkov wavefront not dissimilar to the sonic boom created by aircraft travelling faster than the speed of sound.

This process is visualised in Figure 3.5. The angle with which the cone is created is determined from the local phase velocity of light $c' = c/n$, where n is the refractive index of light, and v the velocity of the charged particle such that

$$\cos\,\theta_C = \frac{1}{\beta \cdot n}, \tag{3.3.2}$$

where $\beta = v/c$. By requiring that the velocity of the particle is larger than the local speed of light, i.e. $v > c' = c/n$ and therefore $\beta > 1/n$, where

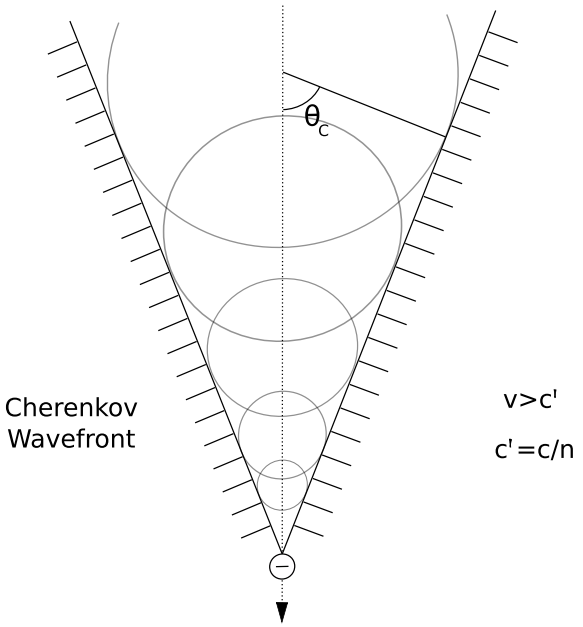


Figure 3.5: As a relativistic charged particle passes through a medium at a velocity greater than the local speed of light, a trail of collapsing polarisation states results in a coherent wavefront of Cherenkov radiation.

$$\beta = v/c = \sqrt{1 - (E_0/E)^2}, \tag{3.3.3}$$

the minimum energy with which Cherenkov radiation can be produced can be determined as

$$E_{thresh} = \frac{m_0 c^2}{\sqrt{1 - n^{-2}}}. \tag{3.3.4}$$

The spectrum of Cherenkov radiation that is produced is governed by the Frank-Tamm Formula, which describes the number of Cherenkov photons emitted as a function of wavelength λ and path length x which can be simplified to

$$\frac{d^2N}{dxd\lambda} \propto \left(1 - \frac{1}{\beta^2 n^2}\right) \frac{1}{\lambda^2}, \tag{3.3.5}$$

Ignoring the refractive index and the speed of the particle for the time being, this will produce a spectrum in the form of $1/\lambda^2$ which results in a preference for photons of shorter wavelengths. In practice, due to atmospheric absorption, the peak Cherenkov radiation occurs at around 300-400 nm. i.e. in the blue/UV range.

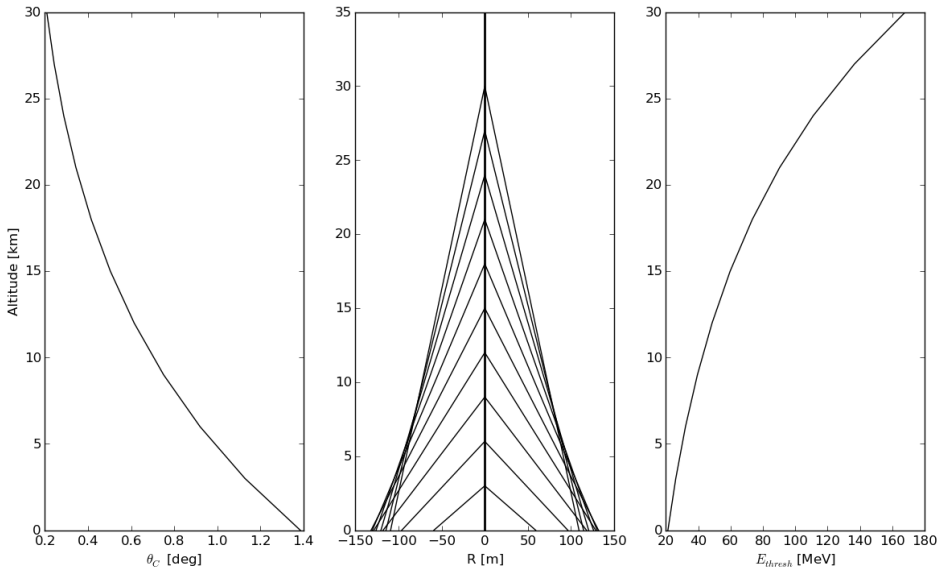


Figure 3.6: Evolution of Cherenkov angle, energy threshold and radius of the resulting Cherenkov light cone at observation level. The evolution of the refractive index within the atmosphere results in a focusing of Cherenkov light into a ring of around 120 m in radius.

Both the Cherenkov emission angle and the threshold energy are a function of the refractive index of air, which in turn is a function of wavelength and altitude (based on the increasing density of the atmosphere with decreasing altitude). Using an atmospheric profile for a site similar to that of current ground-based gamma-ray instruments, the variation with altitude can be seen in Figure 3.6. The change in Cherenkov angle with altitude results in a focusing of the Cherenkov light in a ring around the shower axis of radius ∼120 m, calculated using

$$R = \frac{h - H}{\tan \theta_C}, \tag{3.3.6}$$

where h is the starting altitude and H is the observation level. This distribution of Cherenkov emission at ground level from a 300 GeV gamma-ray, along with a 1 TeV proton, can be seen in Figure 3.7. The spread of light outside of R in the case of the gamma-ray comes predominantly from the Coulomb scattering of the $e^- e^+$ pairs in the shower development. The complex nature of the proton shower results in a disordered distribution of light on the ground with small sub clusters

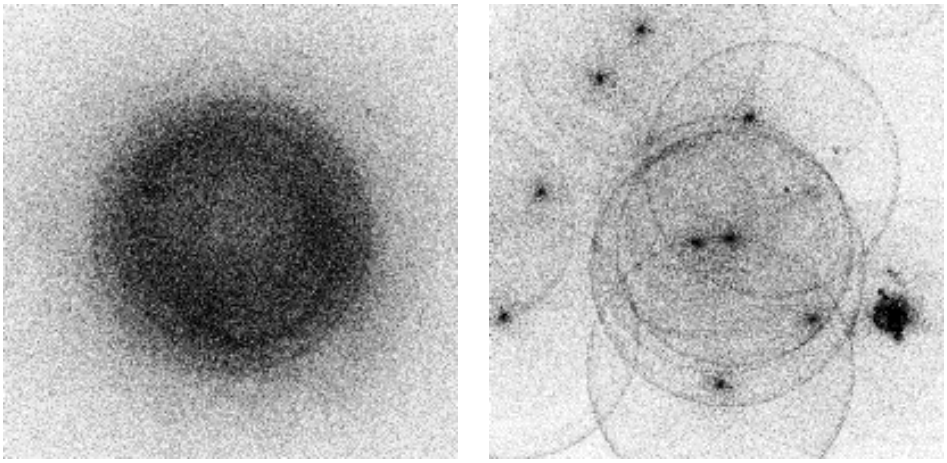


Figure 3.7: Distribution on the ground of Cherenkov light from a 300 GeV gamma-ray (left) and a 1 TeV proton (right) produced by K. Bernlöhr at the MPIK in Heidelberg. The complex nature of a hadronic air shower results in multiple components from hadronic cores, secondary electromagnetic showers and muon rings.

appearing from the separate secondary electromagnetic showers along with rings created by muon showers (See Chapter 5). The difference in these distributions aids the separation of signals from gamma-ray events and background cosmic ray events in ground-based gamma-ray astronomy.

3.3.4 Shower Imaging

We have seen that air showers generated by both the desired signal of gamma-rays and the background of cosmic rays result in a flash of Cherenkov radiation that is observable at ground level. It will now be shown that by using a telescope with a large collection area and fast electronics it is possible to sample the EAS and reconstruct the direction and energy of the primary gamma-ray, and to reject the background cosmic rays. The cameras used in this method require very sensitive photodetectors. Traditionally, these have taken the form of photomultiplier tubes but with the progress of technology (due to a need in scientific instruments and in medical imaging), alternatives such as multianode photomultipliers (MaPMs) and silicon photomultiplier (SiPMs) are currently under investigation; these will be discussed in the next Chapter. If a shower occurs within the field of view of the telescope, the camera will be able to sample the Cherenkov light which, for

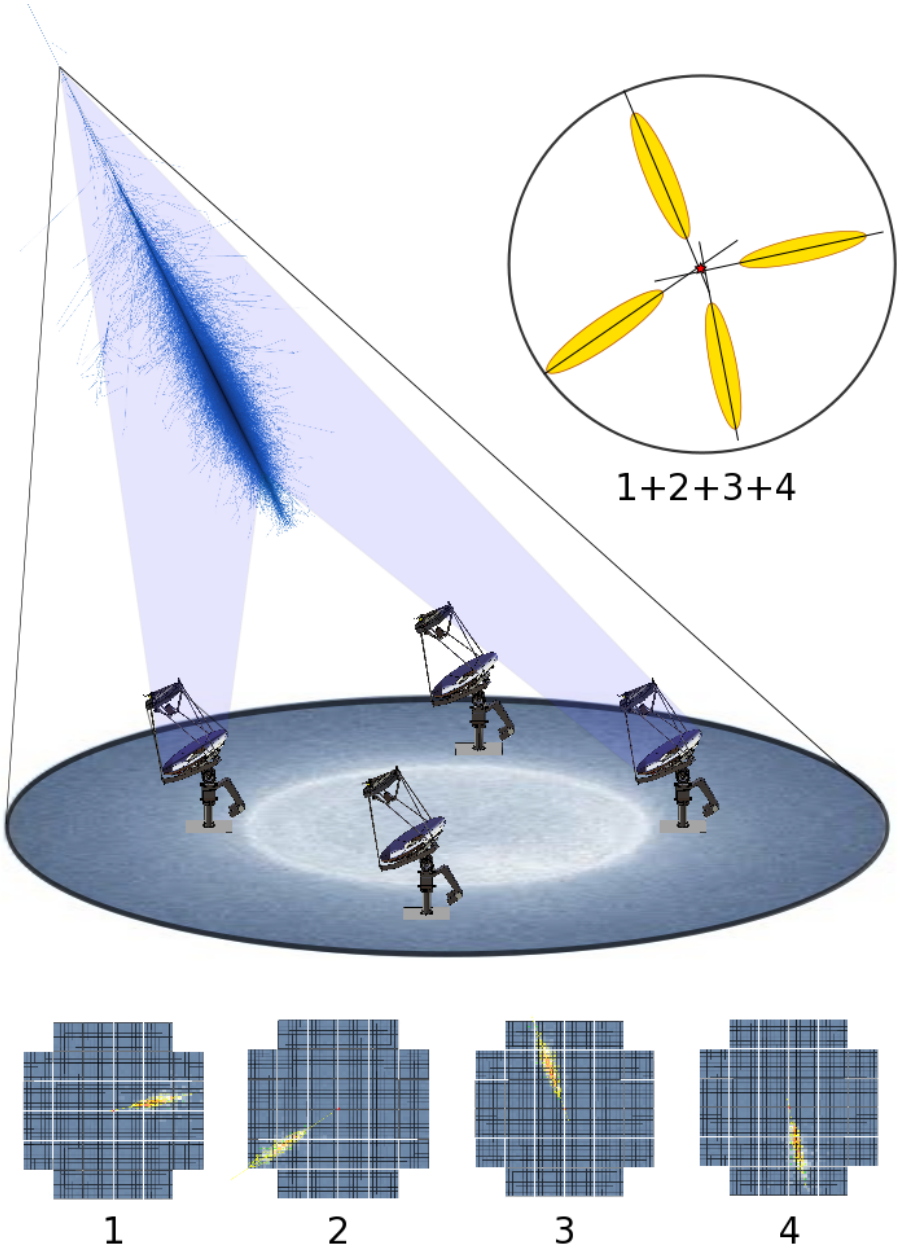


Figure 3.8: Superposition of four camera images observing the same EAS. By using the stereoscopic technique, an improved reconstruction of the shower parameters can be achieved.

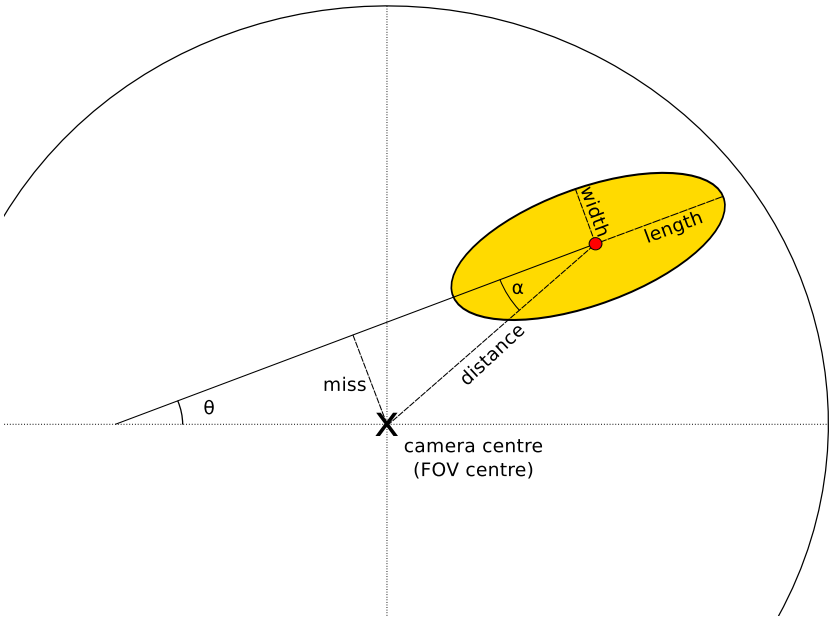


Figure 3.9: Hillas parametrisation of a shower image within a camera modelled by an ellipse.

electromagnetic induced air showers, will be encompassed by a circle if on axis or an ellipse otherwise. See Figure 3.8 for an example.

The ellipse that is imaged within the camera is then parametrised in order to reconstruct the shower and obtain information such as the shower-axis orientation, the core position in the observation plane, the angular dimensions of the shower and the depth of the shower maximum. The most widely used method for this was proposed in 1985 by Hillas [76] in which the image is defined as function of the second moments obtained from a fit of an ellipse to the image (see [66]). The first order moments are the coordinates of the center of the ellipse, defined as

$$\langle x \rangle = \frac{\sum_i I_i x_i}{\sum_i I_i}, \qquad \langle y \rangle = \frac{\sum_i I_i y_i}{\sum_i I_i}, \tag{3.3.7}$$

where x_i and y_i are the pixel coordinates and I_i is the pixel intensity post calibration for each pixel i. The second moments are defined as:

$$\langle x^2 \rangle = \frac{\sum_i I_i x_i^2}{\sum_i I_i}, \qquad \langle y^2 \rangle = \frac{\sum_i I_i y_i^2}{\sum_i I_i}, \qquad \langle xy \rangle = \frac{\sum_i I_i x_i y_i}{\sum_i I_i}. \tag{3.3.8}$$

From these the variance and covariance of each parameter are calculated

$$\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2, \qquad \sigma_y^2 = \langle y^2 \rangle - \langle y \rangle^2, \qquad \sigma_{xy} = \langle xy \rangle - \langle x \rangle \langle y \rangle. \tag{3.3.9}$$

Finally, the following quantities are defined

$$d = \sigma_x^2 - \sigma_y^2, \qquad z = \sqrt{d^2 + 4\sigma_{xy}^2}, \qquad u = 1 + \frac{d}{z}, \qquad v = 2 - u. \tag{3.3.10}$$

From these the Hillas parameters can be derived where a physical representation can be seen in Figure 3.9. These are specified as

- Width: the width of the ellipse is related to the lateral development of the shower
$$W = \sqrt{\frac{\sigma_x^2 + \sigma_y^2 + z}{2}}. \tag{3.3.11}$$
- Length: the length of the ellipse is related to the vertical development of the shower
$$L = \sqrt{\frac{\sigma_{x^2} + \sigma_{y^2} - z}{2}}. \tag{3.3.12}$$
- Distance: the distance between the image centre of gravity (COG) and the centre of the field of view which is related to the distance between the shower core impact point and the telescope
$$D = \sqrt{\langle x \rangle^2 + \langle y \rangle^2}. \tag{3.3.13}$$
- Size: the image size, or the total integrated light content of the shower, which is related to the energy of the shower primary
$$S = \sum_i I_i. \tag{3.3.14}$$
- Miss: the perpendicular distance between the image major axis and the centre of the field of view which is related to the shower orientation
$$M = \sqrt{\frac{1}{3}(u\langle x \rangle^2 + v\langle y \rangle^2) - \frac{2\sigma_{xy}\langle x \rangle \langle y \rangle}{z}}. \tag{3.3.15}$$

- Alpha: the angle between the image major axis and the distance parameter which is correlated to the shower direction

$$\alpha = \arcsin\left(\frac{M}{D}\right). \tag{3.3.16}$$

Originally this technique was developed for use with the Whipple telescope and can be used to reconstruct an event with a single telescope image. However, with all current ground-based gamma-ray telescopes, a minimum of at least 2 telescopes participating in an event trigger is required. This stereoscopic trigger results in a more robust understanding of the EAS by reducing the uncertainty on reconstructed parameters. Again consider Figure 3.8 where four telescopes record an image of the same shower. In this case, for example, the images can be combined and the intersection of the major axis of each ellipse reveals the source direction.

The use of Hillas parameters is not only useful in reconstructing the shower geometry, it can also be used for the rejection of cosmic ray background. Recall from Figure 3.7 the difference in the distribution of the Cherenkov light pool from photons and hadrons at ground level. From this it should be clear that the images produced by hadronic EAS will be largely disordered compared to the clean ellipse of an electromagnetic EAS. Therefore it is possible to compare the measured image width and length, for example, to an expected value based on the reconstructed impact distance and image size. These expected values are obtained a priori by creating large scale Monte Carlo simulations of the system (see Chapter 4) which provide a lookup table of parameter distributions. In order to combine information from multiple telescopes the two normalised parameters, scaled width (SW) and scaled length (SL) are defined such that

$$SW = \frac{W - \langle W \rangle}{\sigma_W}, \qquad SL = \frac{L - \langle L \rangle}{\sigma_L}, \tag{3.3.17}$$

where W (L) is the width (length) of the image, $\langle W \rangle$ ($\langle L \rangle$) and σ_W (σ_L) are the expected value and spread obtained from simulation look up tables and are a function of size and reconstructed impact distance. This is then weighted by the number of telescopes, *ntel*, participating in the reconstruction to produce the mean scaled reduced width and mean scaled reduced length

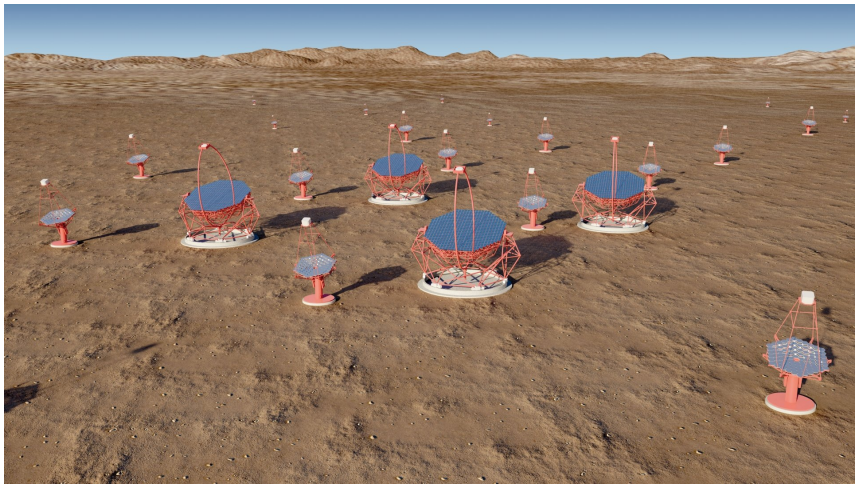


Figure 3.10: Artist's impression of an example array of CTA telescopes, consisting of 4 central LSTs surrounded by a number of MSTs with SSTs present in the distance. Figure created for CTA by DESY/Milde Science Comm./Exozet

$$MSCRW = \frac{\sum_{ntel} SW}{\sqrt{ntel}}, \qquad MSCRL = \frac{\sum_{ntel} SL}{\sqrt{ntel}}. \tag{3.3.18}$$

Based on the expected distribution of gamma-ray and cosmic ray images, cuts can be applied to each observed event. This is known as the standard cuts method and is widely used to reject cosmic ray background whilst maintaining a good gamma-ray efficiency [16]. Inevitably, no background rejection method is perfect and therefore more complex methods, but overall still requiring the Hillas parameters, are used and are currently being further developed (See Section 4.5 for an example). The current main concentration is towards creating a more robust pipeline for the next generation of ground-based gamma-ray instruments, namely CTA.

3.4 Cherenkov Telescope Array

With current ground-based gamma-ray instruments still being considered as experiments (see H.E.S.S. [78], VERITAS [121] and MAGIC [86]), operated mainly by the groups that designed them, CTA will represent a large step in the relatively new field of gamma-ray astronomy. Operating as an open observatory, driven largely by guest observation proposals, CTA will ensure that it becomes an essential asset to the wider astrophysical community. While still building on the technology of the

previous generation, it is largely the increase in scale that will result in a great leap forward, moving from the current largest array of 5 telescopes to two arrays of 50-100 telescopes covering both hemispheres. See Figure 3.10 for a conceptual image of CTA.

The main aims for CTA are to increase the maximum sensitivity to 10 times the current achievable limit. In the core energy range of some 100 GeV to several TeV this corresponds to a level 10^3 times below the flux of the brightest known steady source at VHE energies, the Crab. Along with this it is foreseen that the angular resolution will reach levels of 2 arc minutes at around 1 TeV and an energy resolution <10% above 1TeV [3]. In addition, the aim is to achieve these improvements over 4 orders of magnitude in energy, from a few tens of GeV to hundreds of TeV, allowing unprecedented studies of the physical processes occurring in VHE sources to be performed. In order to facilitate this large energy range, CTA will consist of three different sized telescopes designed to have optimal performance at different energy ranges:

- **Low Energy:** For primary gamma-rays with energies from 20 GeV to around 200 GeV, the induced EAS are frequent, however the amount of Cherenkov light observable at ground level is low ($\sim$1 photo electron (p.e.) m^{-2}). Therefore the optimal design for this energy range would be a small number of telescopes with large collection areas in order to focus enough of the Cherenkov light to image the lowest energy showers. For CTA this will come in the form of 3-4 large size telescopes (LSTs) with $\sim$23 m diameter mirrors.

- **Medium Energy:** This represents the energy range at which most current IACTs are sensitive, 100 GeV to 10 TeV, therefore the design of the medium sized telescopes (MSTs) will be somewhat similar to current designs like H.E.S.S. and VERITAS but covering a greater area with $\sim$12 m diameter telescopes, to achieve the goal CTA sensitivity. There is also an alternative design being investigated that uses a two mirror system, which will be described in Section 3.5.

- **High Energy:** At high energies, over a few TeV, primary gamma-rays are

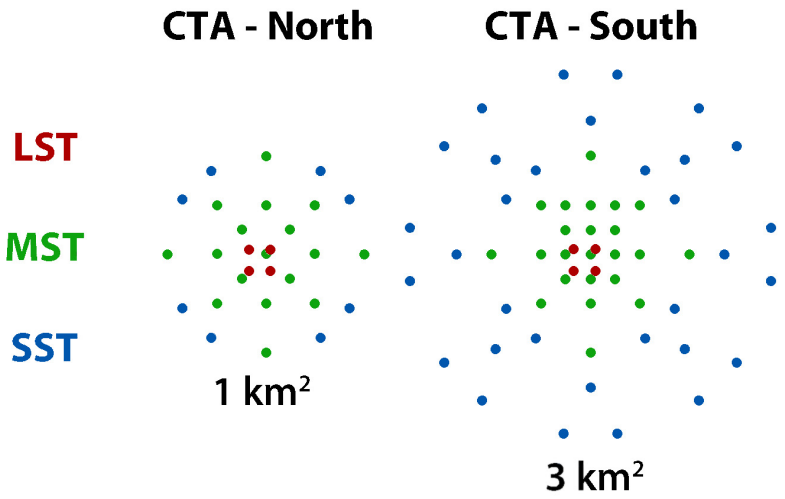


Figure 3.11: Early example layouts for the northern and southern array. Optimisation of the layout is in the final stages of completion. Credit: the CTA consortium.

rare but due to this great energy the resulting EAS are rich in Cherenkov radiation ($\sim 10^3$ p.e. m^{-2}). An optimal design would therefore be in contrast to the LST, with numerous small size telescopes (SSTs), with ∼4 m diameter mirrors covering as large an area as possible while still retaining a reasonable number of telescopes within the stereoscopic trigger. To facilitate this, it is also desirable for the SSTs to have a large field of view.

The arrangement of these telescopes within an array has recently been finalised (Kashiwa meeting May 2016). An example of early layout proposals can be seen in Figure 3.11 and the resulting differential flux sensitivity (point source sensitivity for 5σ detection within a given amount of time), angular resolution (68% containment radius) and energy resolution (spread in reconstructed energy) for the southern array can be seen in Figure 3.12. For more detail on the CTA project as a whole, see [3]. For the remainder of this work, the focus will be on one of the proposed solutions for the SST component of the array, the Gamma Cherenkov Telescope (GCT).

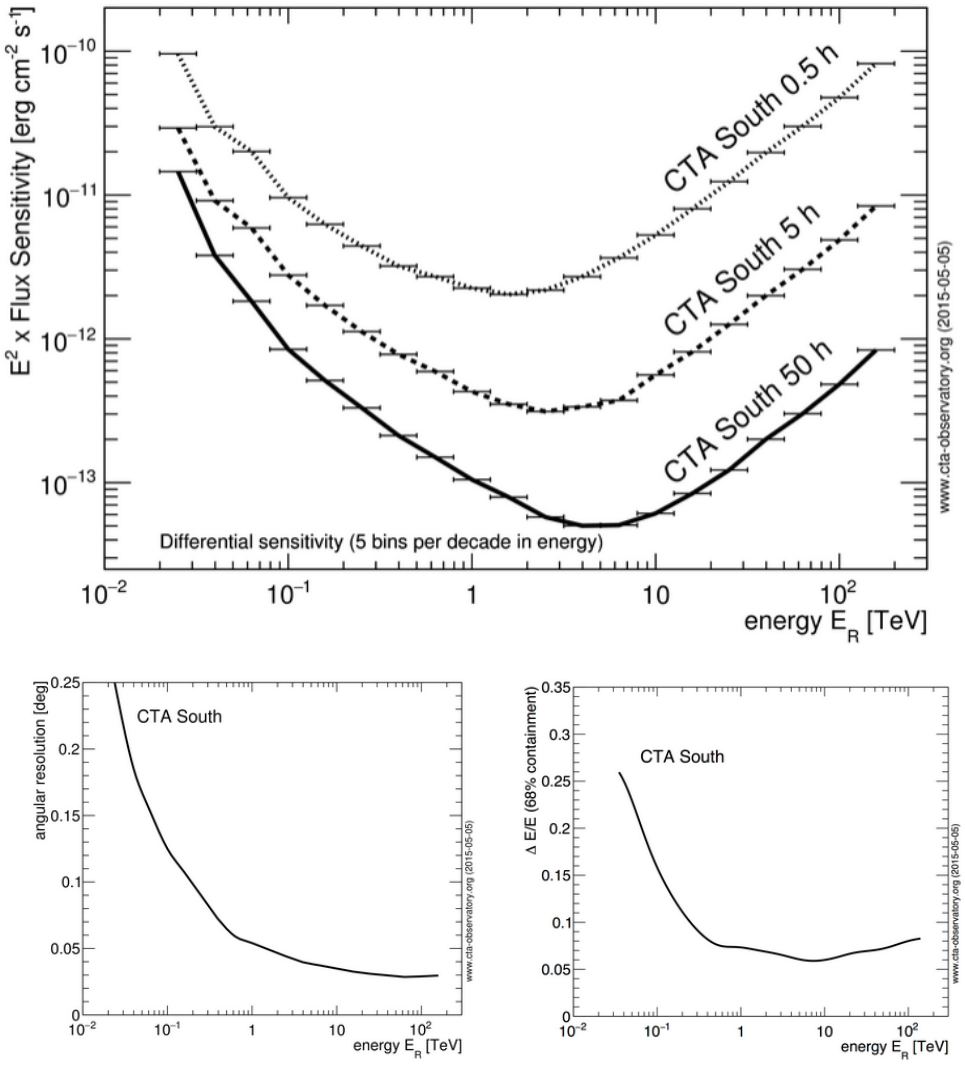


Figure 3.12: Current expected CTA performance for the southern array. Credit: the CTA consortium.

3.5 GCT: High energy SST component of CTA

The EAS produced by the highest energy gamma-rays (above some some 10s of TeV) result in a strong signal of Cherenkov radiation compared to those of lower energies, but due the generally steep energy spectra of astrophysical sources, the challenge is to observe enough events to achieve any statistically significant information. As was mentioned in the previous Chapter, CTA aims to cover a wide energy range by using three different size telescopes, with the small size telescope (SST) covering energies of $\sim$1 TeV to several 100's of TeV. In order to meet the goal sensitivity in this energy range, the SST component of CTA must cover a very large area. With this in mind, a focus has been placed on producing a design that will reduce costs while also achieving a large field of view enabling the required effective area to be achieved with fewer telescopes (which incidentally will also help with surveys and studies of extended objects). This design challenge has necessitated the move to innovative technologies such as the use of silicon detectors which are rapidly increasing in performance and decreasing in cost and have been shown to be successful in the FACT telescope [19]. Additionally, it is possible to implement more complex optics such as introducing a secondary mirror to the system which allows for a larger field of view.

Traditionally, IACTs have been based on Davies-Cotton (DC) or parabolic single reflector designs, comprising of a single large segmented primary reflector with the camera located in the focal plane. If a large field of view is required, this design becomes somewhat problematic, requiring a larger reflector, longer focal length and bigger camera. This places increasing constraints on the mounting structure of the optical system in order to prevent the occurrence of significant optical aberration, resulting in poor image quality off axis. One solution to this is to include a secondary mirror in order to demagnify the air shower image. This was in fact first suggested for optical telescopes in 1905 by Karl Schwarzschild and later improved in 1926 by Andre Couder to include a curved focal plane [119]. However, the design remained unchanged and unbuilt until it was seen as a solution to the SST large field of view problem. The Schwarzschild-Couder (SC) geometry allows for a reduction in the scale of the focal plane and the physical pixel size while retaining a good

angular resolution out to large field angles. The small pixel size also allows the use of commercially available multi-pixel photodetectors such as multi-anode photomultipliers (MaPMs) and silicon photomultiplier (SiPMs), leading to a relatively low cost camera.

For CTA, there are currently three separate designs for the SSTs, one using a traditional DC design, the SST-1M [96], and two with SC optics, the Italian ASTRI [100] and the Gamma Cherenkov Telescope (GCT) comprising a French[2] SST-GATE (Small Size Telescope GAmma-ray Telescope Elements) structure and the CHEC (Compact High Energy Camera) camera worked on by a consortium of six countries: the UK, Germany, US, Japan, Australia and the Netherlands. It is the GCT that will be the main focus of this work, with the rest of this Chapter dedicated to a description of the structure and camera. In Section 3.5.1 the SST-GATE structure will be described, followed by the CHEC camera in Section 3.5.2. Finally, with CTA moving rapidly through its prototyping phase, the first GCT built in Paris, along with initial results, will be described in Section 3.5.3.

3.5.1 Structure: SST-GATE

Although the first design for a SC prototype SST structure for CTA was developed at the University of Durham, it was further optimised after being taken on by the SST-GATE group. The current design was created with the primary idea of optimising the point spread function (PSF) on axis, allowing an increase at larger field angles, but still within the pixel size of the camera. A secondary goal was to reduce the mass and cost of the structure while allowing easy mounting and maintenance. As a result, GCT is visually very different than the second SC-SST design, ASTRI, which was optimised to obtain a pointing precision of 7 arcseconds without the need for further calibration, resulting in a structure that is over twice the weight of GCT [62]. A diagram of GCT can be seen in Figure 3.13 in which the main elements have been separated into the alt-azimuth structure, the optical support structure

[2]With contributions from the UK and Germany to investigate mirror solutions and the Netherlands looking at pointing calibration.

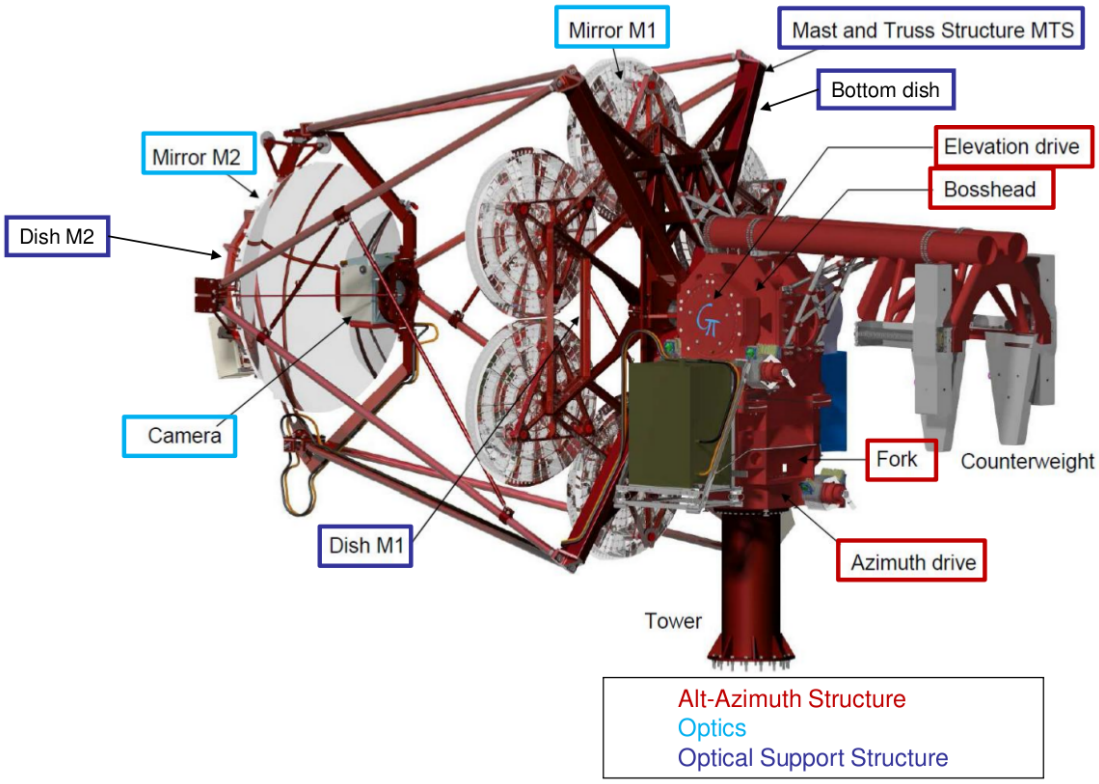


Figure 3.13: Figure showing the mechanical structure of GCT (SST-GATE), highlighting the separate components which in turn are grouped into the alt-azimuth structure, the optical support structure and the optics. Credit: Observatoire de Paris.

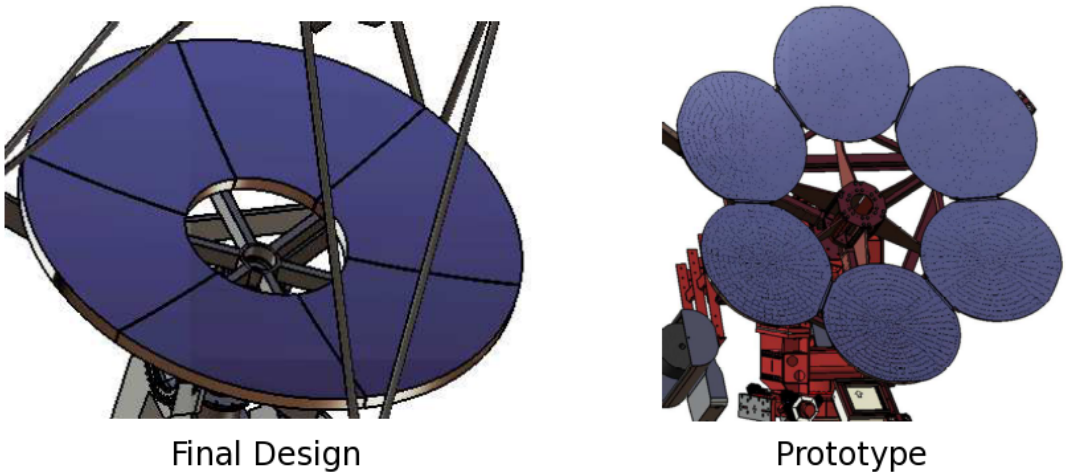


Figure 3.14: The planned primary mirror layout (left) and the prototype mirror layout (right). Downscaled mirrors were chosen for the prototype due to the ease of manufacturing and coating which in turn brought down costs. Image taken from [63], credit: Observatoire de Paris.

and the optics. The alt-azimuth structure is similar in design to that used with the Whipple and VERITAS telescopes, for a more detailed description, see [60] as it is the optics that will become relevant later in this work. The mast and truss structure, designed in the Serrurier configuration of 8 steel tubes, has been created to be stiff enough to hold the weight of the optical structure but still remaining as light as possible. To this the mirror support dishes (dish M1 and dish M2) are attached on which the telescope mirrors are mounted. The dish M1 will be constructed in such a way that allows easy mounting and maintenance of the telescope mirrors, with the ability to rotate around the optical axis and provide easy access at ground level.

The mirror configuration that can be seen in Figure 3.13 represents the current mirror design for the GCT prototype which will be discussed in Section 3.5.3. For the final telescope, the primary reflector will be tessellated into 6 petals, creating a 4 m diameter circular dish with a 1.3 m diameter hole in the centre where light is blocked by the secondary reflector; this can be seen in Figure 3.14. The secondary will also consist of 6 petals, but unlike the primary, they will be bolted together to form a monolithic mirror. Each mirror (the 6 primary segments and the monolithic secondary) will then be aligned with 3 actuators. While work is still ongoing into investigating mirror solutions for CTA, the current design for GCT is to machine bulk aluminium samples which are then polished to reach a low roughness (less than 0.02 μm) and then coated with either aluminium, nickel or a dielectric material in order to reach the required reflectivity[3]. The two reflectors are aspherical, with substantial deviations from the closest spherical shape, separated by a distance of 3.56 m and then a further 0.511 m between the secondary and the camera. The resulting focal length from this geometry is 2.283 m and a plate scale on the camera of 0.025 degree/mm. (For more information about the current telescope mirrors see [63]).

[3]It is worth noting, that while it is desirable to have high quality mirrors, they do not have to meet the same optical standards as those for optical telescopes. A greater emphasis must be placed on obtaining a good reflectance in the wavelength range appropriate for Cherenkov light ($\sim$300 to $\sim$500 nm), while also ensuring a robust surface which can survive long exposure to the elements, as IACTs are not generally protected by domes.

3.5.2 Camera: CHEC

The CHEC camera, now in a joint collaboration with the SST-GATE telescope group, is designed to be compatible with both the SC SST telescopes. The main ethos was to create a camera that was light, compact and cost effective while still providing a high quality output [56]. Due to the reduced plate scale provided by the SC optics, it was possible to investigate two possible "off the shelf" compact detectors. In a large part due to the advances of medical imaging techniques, commercially developed multi-anode photomoltipliers (MaPMs) and silicon photomoltipliers (SiPMs) have become attractive options for use in Cherenkov cameras. Therefore it was decided that these two solutions would be investigated, the first using MaPMs, from now on referred to as CHEC-M (or GCTM if describing the whole telescope), which has been built and is undergoing tests and commissioning, the second with SiPMs, referred to as CHEC-S (or GCTS) of which a prototype is now being constructed. Many aspects of each camera are common. In the following, an overview of CHEC design (see Figure 3.15) and functionality will be given which will become important in Chapter 4.

- **Mechanics** The mechanics of the CHEC camera covers the focal plane, outer casing, internal racks for the electronics, the thermal control unit and an interface plate at the rear for mounting to the telescope. These components can be seen in Figure 3.16. The focal plane plate, in which the photodetectors are placed, is made of a machined anodised aluminium block with a 1 m radius of curvature. Once fully assembled, the casing will be sealed to prevent ingress of water or dirt, aided also by the carbon fibre lid which will additionally protect the sensitive photodetectors. The temperature within the camera is regulated by a thermal unit mounted on the side of the camera which consists of fans coupled to a water-cooled heat sink; this will keep the internal air temperature to below 45°.

- **Detectors** CHEC will consist of 32 photodetector modules each with 8×8 pixels, making up a total of 2048 pixels. The MaPM can be visualised as many PMTs contained within one housing with multiple anodes at the rear.

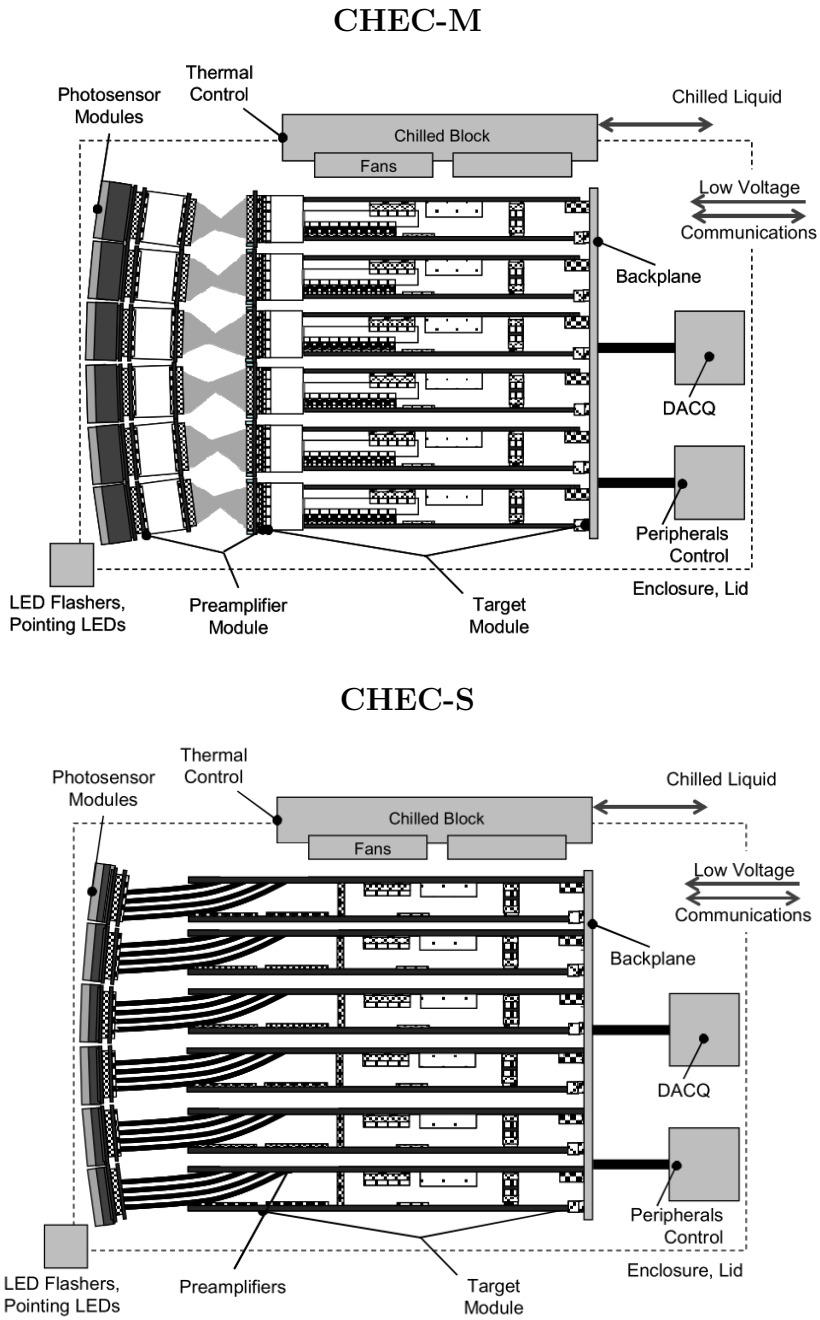


Figure 3.15: Side view of the internal structure of the cameras showing the important components which are described in the text. Image from [62].

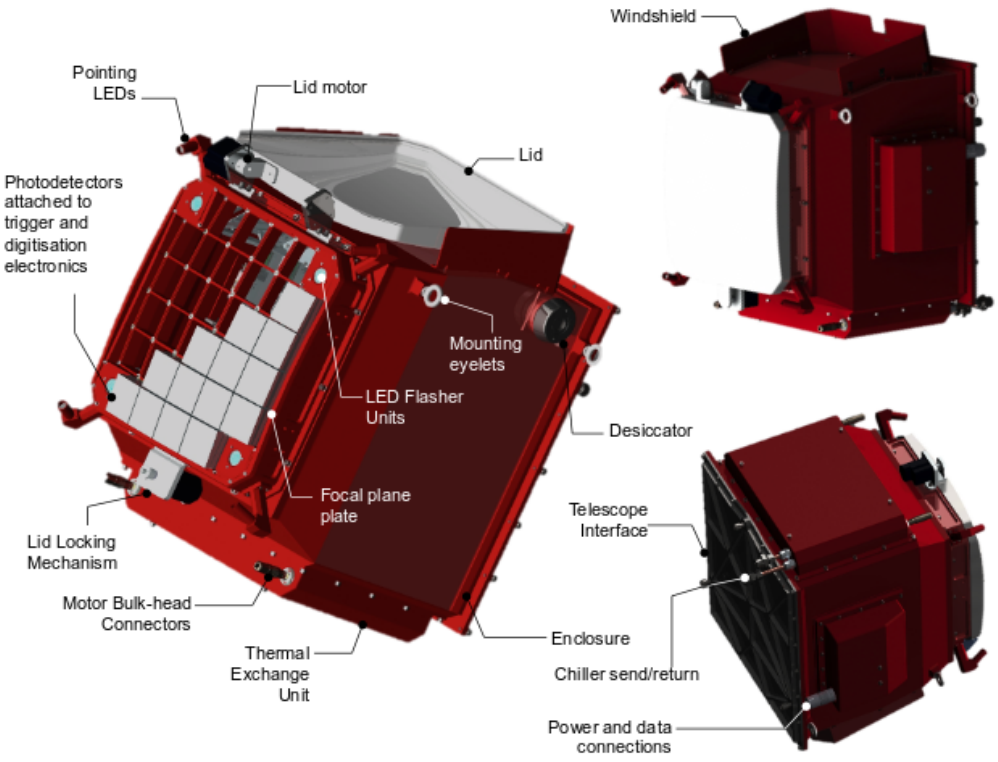


Figure 3.16: 3D model of CHEC-M highlighting the individual components. Image from [56].

Each module is $\sim 52 \times 52$ mm and $\sim$27 mm deep, which, due to the curvature of the focal plane, results in a gap between modules of around 5 mm. The operational gain of the photomultipliers is expected to be 8×10^4 ($\pm \sim 25\%$ between pixels) providing a dynamic range of 1000s of p.e.. In the case of the SiPMs, the modules measure 51.4 mm across but are much thinner, therefore the resulting gap between modules is much smaller ($\sim$1 mm). Work is currently under way to fully characterise the SiPMs.

- **Pre-Amplifiers** Once a signal has been generated within the photodetector module, it is passed to the pre-amplifier via micro coaxial cables that remove the radius of curvature of the focal plane. The main purpose of the pre-amplifier if to perform fast amplification and to shape the detector module signal into a pulse of 5.5-10.5 ns FWHM, allowing for optimal coincidence time windows for triggering. In the case of the CHEC-S the pre-amplifier also supplies each silicon pixel with an operational bias voltage.

- **Target Modules** The amplified signal is then digitised within the Target modules [115], where each channel is digitised at 1GSa/s over a programmable 96 ns window. This is also where the first level camera trigger is performed. An analogue sum of 4 neighbouring pixels is taken and must be above a certain threshold, here referred to as the discriminator threshold; These are defined as trigger pixels. The Target module also routes the HV required by the MaPM modules.

- **Backplane** The signals from each Target module are routed to a large single PCB known as the backplane. Here, the second level trigger is evaluated in which 2 or more neighbouring trigger pixels must satisfy the first level trigger, within a programmable coincidence window. The precise timing needed in this process is provided to each camera via a White Rabbit timing network [109]. The backplane also sends controls to the rest of the camera (via a peripherals board) to control the lid and the calibration units among other things.

- **DACQ** The data acquisition system (DACQ or DAQ) receives the data from the target modules routed via the backplane. From here the data, along with

Figure 3.17: Images of the prototype telescope in Paris during the inauguration, photo credit: A. Okumura.

control and monitoring information, are routed to and from the camera.

- **Calibration** At each of the four corners of the camera, there is a calibration unit, each containing 10 UV LEDS which are used to produce fast light pulses to uniformly illuminate the camera. These UV *flashers* will be used to measure the single p.e. response at the start and end of observations as well as to perform flat fielding data during operation [42].

3.5.3 Prototype

In order to validate the design for GCT, over the last few years construction of prototypes has been under way on both the SST-GATE structure at the Observatoire de Paris and the CHEC-M camera at the University of Leicester. Most recently, during the first on-site tests of the telescope structure and camera in November 2016

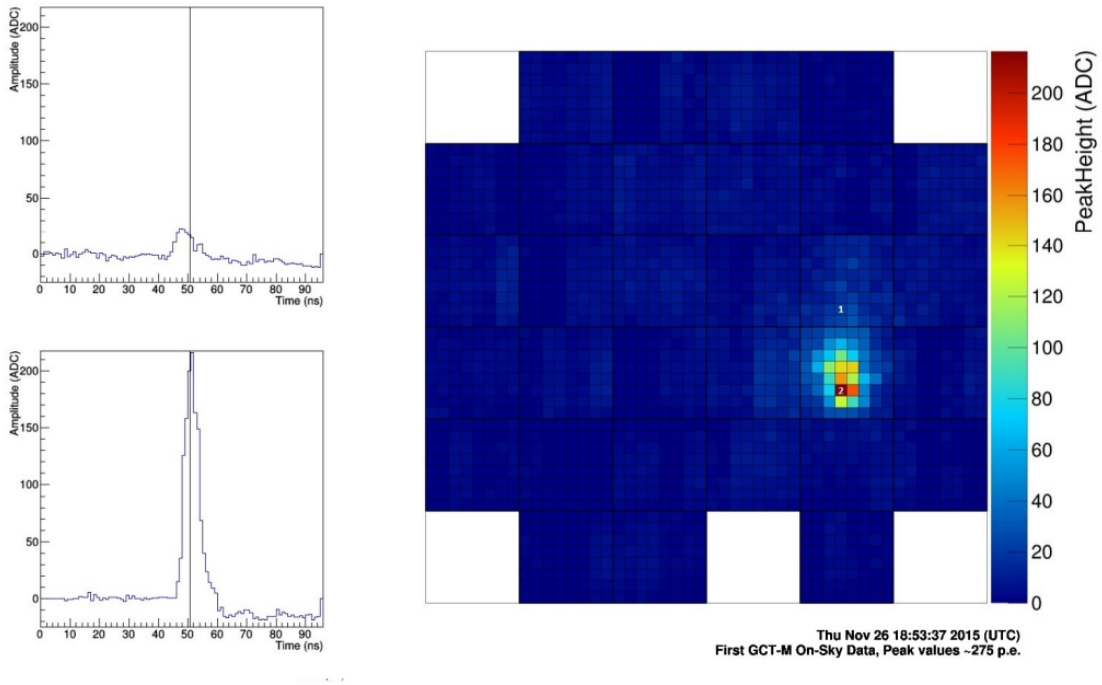
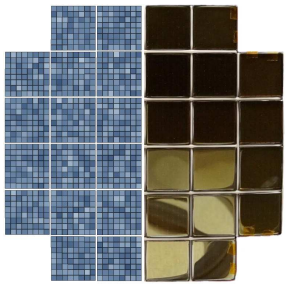


Figure 3.18: One of the shower image taken during the first run with the prototype. The left panels show the measured Cherenkov pulse shape within two example pixels. Due to a technical issue there is one non functional module (bottom). Credit: The GCT sub consortium.

(see Figure 3.17), the first air shower imaged by a CTA prototype was recorded and can be seen in Figure 3.18. This represented a major step for GCT and for CTA along with highlighting the importance of understanding the system. To this end, in the next Chapter work focusing on fully representing the telescope within Monte Carlo simulation will be presented.

Chapter 4

Monte Carlo Simulations for GCT and CTA



In this Chapter the outline of the Monte Carlo simulations performed for the GCT will be presented. The initial focus will be on building a telescope in the simulation software framework of SIM_TELARRAY, which for GCTS, due to similarities in design, received contributions from the ASTRI project. The analysis performed on an array of telescopes is performed with a neural network based background rejection algorithm developed in Durham, where a full description can be found in [105].

4.1 Introduction

For complex systems for which a full analytical solution would be impractical, i.e. in systems with a large number of degrees of freedom, an alternative approach can be be adopted. By drawing numbers randomly from an assumed or measured probability distribution, an approximate solution can be derived; this is known as the Monte Carlo (MC) method. From the discussion in Chapter 2, it should be clear that one such complex system is the development of EAS. While a simplified explanation based on the Heitler model for the development of an electromagnetic air shower was given, in practice the calculation to arrive at a realistic representation is far more difficult. This requires accurate modelling of the scattering of shower particles, deflections within the magnetic field, ionisation loss, mean free path of particles based on the nucleon cross sections, strong interactions, particle decays and so on. In addition to this, to be able to derive any meaningful results when considering the performance of an array of IACTs for example, a large data set of showers would be needed. Therefore, even when using the MC method, the simulation of EAS is a computationally demanding process.

With the aim of quantifying the possible performance of CTA, several large scale MC simulations have been performed. Each production of simulated data has had a slightly different desired goal: Production 1 (prod 1) aimed to reveal an initial estimate of the capability of CTA; prod 2 worked towards evaluating candidate sites for the northern and southern arrays along with considering alternative telescope layouts; prod 3 was used to settle on the final telescope positions in the preferred observatory locations in Chile and La Palma.

In addition to estimating the potential of a planned array, MC simulations also provide an insight into the performance of current telescopes. By ensuring an accurate representation of each telescope, any degradation due general wear and tear can be monitored (see for example the following Chapter on muon calibration). An overview of the simulation tools used in previous experiments and applied in this work will be given in Sections 4.2 and 4.3 followed by the analysis software in Sections 4.4 and 4.5. With a strong desire to keep the modelled telescope system as close to current knowledge from measurements, an overview of the most important

configuration parameters which have been worked on between prod 2 and prod 3 will be presented in Section 4.6. This will be followed by what is referred to as "low level" evaluation in Section 4.7, representing the performance indicators that can be obtained without a major production of simulated data. In Sections 4.8 and 4.9 more in-depth performance estimators will be derived with the aid of a small production of EAS data for an array of small size telescopes. A summary and discussion of further possible work will be given in Section 4.10.

4.2 Simulating Air Showers: CORSIKA

The first, and most computationally demanding, stage of simulations for ground-based gamma-ray astronomy is in the production of data from EAS by which the primary gamma-rays or cosmic rays are observed. This is performed with the CORSIKA (COsmic Ray SImulations for KAscade) software [74]. The CORSIKA code was originally developed for the KASCADE experiment [21] which set out to understand the makeup of ultra-high energy air showers to aid in the overall understanding of the physics of particle interactions. Now widely used and considered a robust code, it takes into account up-to-date interaction models in order to make precise predictions of EAS [102] (see Chapter 3 for more detail about air shower development). The Cherenkov light produced within the showers is tracked down to the observation level, with the photons grouped into "bunches" in order to reduce computer memory requirements.

An important addition to CORSIKA for the field of ground based gamma-ray astronomy is the IACT option that was developed by Bernlöhr [32]. In simulations of EAS, large amounts of data can be produced if the arrival point of every photon bunch is recorded. Increasing the number of photons grouped within the bunches leads to unusual artefacts, therefore an alternative solution was derived. Instead of recording all photon arrival positions, fiducial spheres encompassing each telescope position are defined. By doing this, only photon bunches that intersect with the shadow of the sphere are recorded. This can be seen in Figure 4.1 in which examples are given.

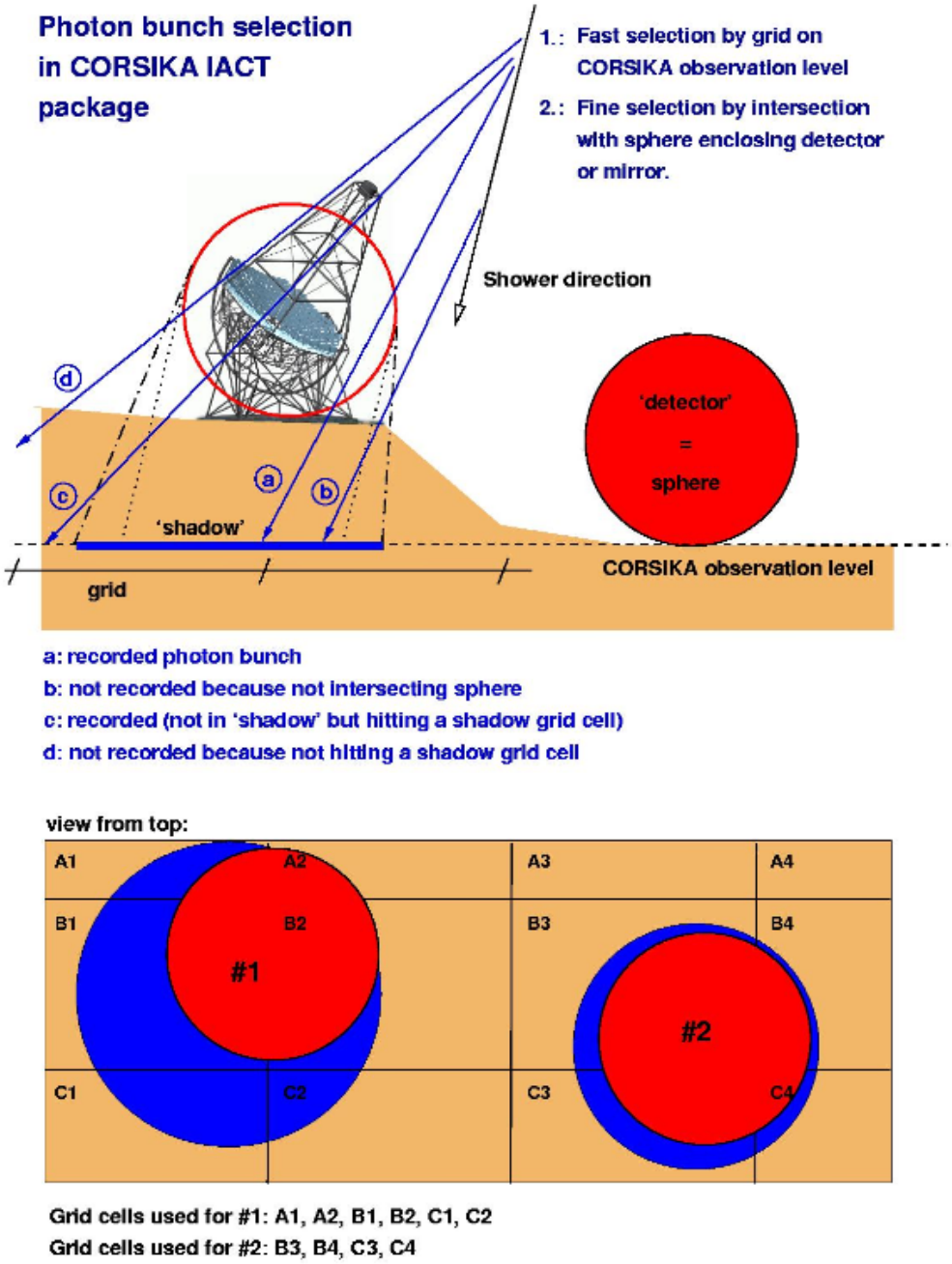


Figure 4.1: The definition of fiducial spheres in CORSIKA using the IACT option. Instead of all photon bunches being recorded within a regular grid at observation level, only those that intersect with the sphere that encloses the whole reflector are recorded. This helps to reduce the amount of data that is stored on disk. Image from [32].

As the atmosphere can in effect be considered as part of the detector for IACTs, it should therefore be apparent that a detailed model of the atmosphere must be made. One of the main inputs that is required for CORSIKA is the atmospheric profile which defines the evolution of the density, refractive index, temperature and pressure as a function of altitude. These all have an effect on the mean free path of particles, and therefore the shower development, along with the yield of Cherenkov light. While Chile has been selected as the main candidate for the southern observatory, the data used in this Chapter were produced prior to that decision[1]. The atmospheric profile was chosen to match the Aar site in Namibia which closely resembles the conditions for H.E.S.S.. While the atmospheric profiles for Aar and Chile are nearly indistinguishable, there are differences in the magnetic field strength (which will cause deflection of the shower particles) and the observation level, 1640 m a.s.l. and 2150 m a.s.l. for Aar and Chile respectively, which will lead to different parts of the shower being imaged. For example, lower altitudes will result in larger shower lateral profiles, increasing the area over which they can be observed but also decreasing the light density. Therefore in this Chapter, especially in Sections 4.8 and 4.9, results should be treated as relative performances. Absolute values may be indicative of the actual performance but should not be taken as final values.

4.3 Optics and Electronics: sim_telarray

Once the light from the EAS has been recorded within the IACT fiducial sphere, the Cherenkov signal must be traced through the telescope structure, optics and camera electronics. This is all performed with the software package SIM_TELARRAY [32]. Originally developed for the HEGRA system [55], this allows for the optical structure, camera configuration, trigger conditions and camera readout for each telescope in an array to be configured separately. These parameters for both GCTS and GCTM will be discussed in detail in Section 4.6, which will also provide a general overview of the simulation process.

[1] An agreement to create this data set was made at the MC-SST meeting in Turin, October 2014.

4.4 Simple Analysis and Processing, read_cta

While SIM_TELARRAY is capable of performing simple image analysis together with simulating the telescope system, more dedicated sets of codes are generally used in order to obtain the event reconstruction. The code used here is READ_CTA and has previously been used to analyse H.E.S.S. data (hence it is often referred to as READ_HESS).

As the output of SIM_TELARRAY essentially consists of waveforms for each camera pixel, the first process that READ_CTA must perform is a calibration of the waveforms (removing NSB contributions and effectively flat fielding the camera) followed by waveform integration in order to obtain a total signal for each pixel. The method used for waveform integration can have an effect on how well the charge is reconstructed and will be discussed in Section 4.7.2. Having obtained an event image, READ_CTA then performs image reconstruction and initial cuts. The image reconstruction uses the Hillas parametrisation as described in Section 3.3.4 and if stereo reconstruction is available, the shower parameters such as energy, direction and impact point are calculated.

4.5 Advanced Analysis, TMVA

To investigate the high level performance indicators for GCT, an analysis pipeline that incorporates the multivariate software package TMVA [79] has been used. The development of this code was carried out previously at Durham University by C. Rulten and a full description goes beyond the scope of this work, therefore only a summary will be presented. For a full description see [105].

As with any analysis, the goal is to obtain the best signal-to-noise ratio possible with high efficiency. In the case of ground-based gamma-ray astronomy, the main background is made up of cosmic ray showers. In Section 3.3 the differences between the images produced by each type of primary, along with the resulting Hillas parametrisation, was presented. The Hillas method is considered effective, but there is always room for improvement and for CTA to reach its full potential, more advanced methods will be needed. One branch of possible improvements comes

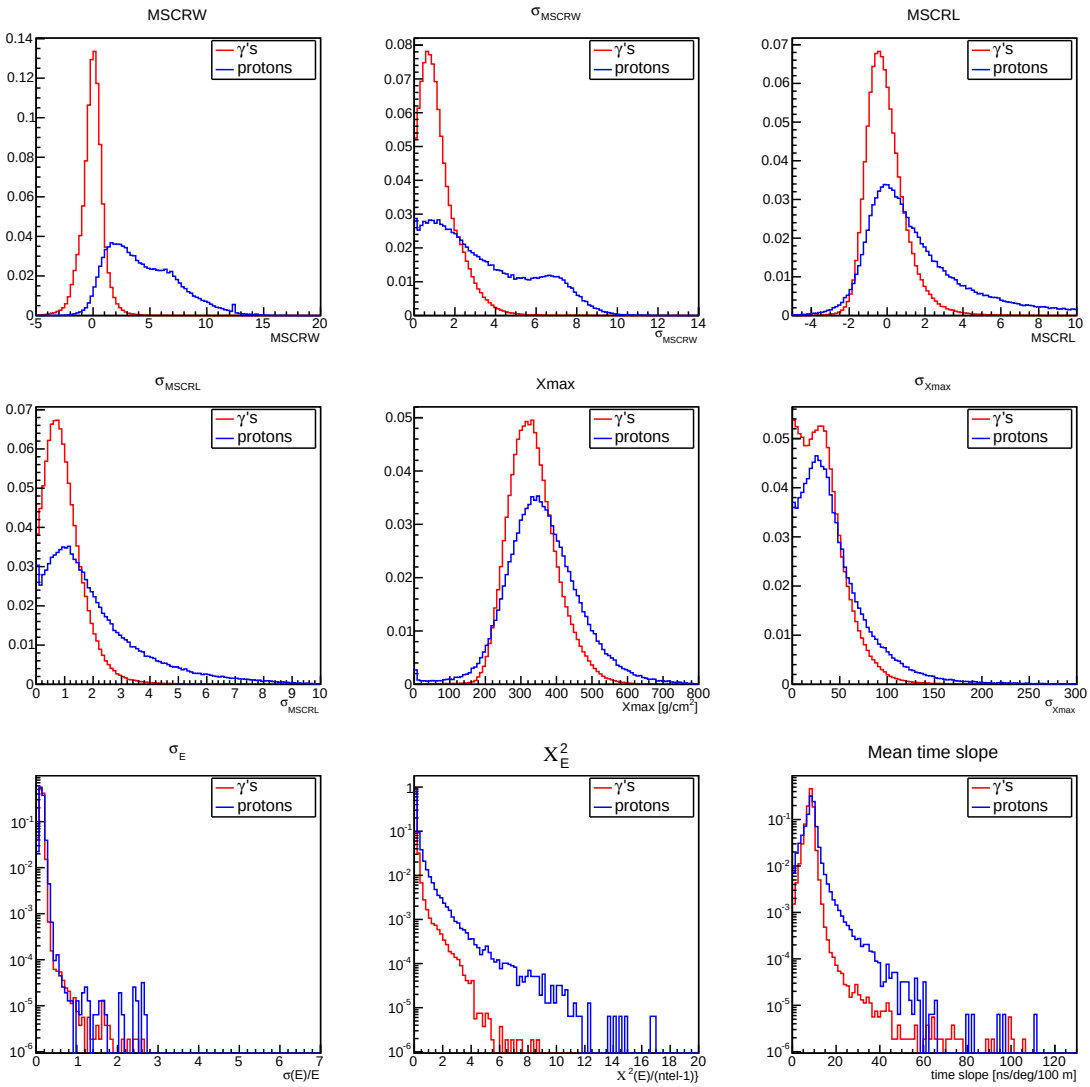


Figure 4.2: The normalised individual training parameters that are used in the TMVA algorithm to produce a single response cut parameter. These are the mean scaled reduced width (MSCRW), the mean scaled reduced length (MSCRL), the spread on those values ($\sigma_{\mathrm{MSCRW}}, \sigma_{\mathrm{MSCRL}}$), the shower depth maximum and its spread ($X_{\max}$, $\sigma_{X_{\max}}$), the spread on the reconstructed energy as well as the fit quality (σ_E/E, $\chi^2_E/(ntel-1)$) and finally the mean time slope. The distributions are shown for both gamma-ray (blue) and proton showers (red) with the largest difference being seen in the MSCRW and MSCRL.

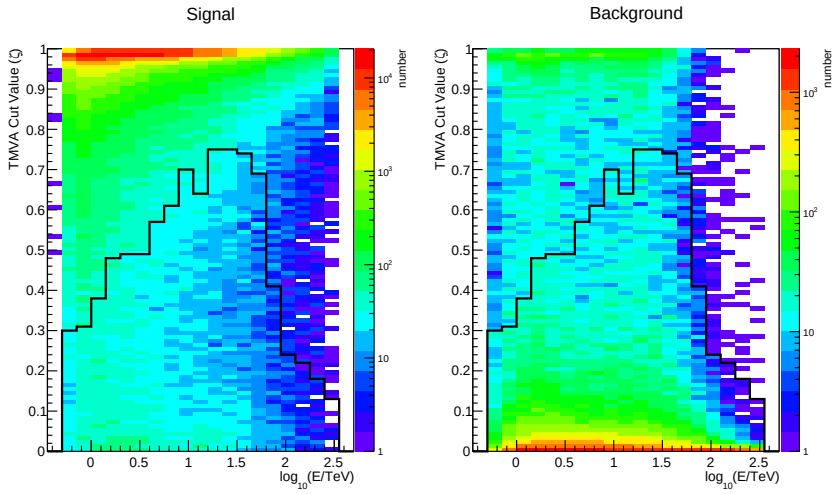


Figure 4.3: The resulting response perimeter (ζ) derived with the neural network shown against the test sample of gamma-ray (left) and proton (right) showers where the z scale shows the density of events. The black line shows the chosen energy dependent cut parameter and it can be seen that this method is effective at using the input parameters seen in Figure 4.2 to separate the signal and background.

from the use of a multivariate approach. This method works by combining all distinguishing factors that separate gamma-ray and cosmic ray showers, such as the image width and length, into a single response parameter. This effectively can be considered as a description of the shower images' gamma-ray "likeness", allowing showers that look like gamma-rays to be kept and rejecting those that don't. Here, a machine learning Multi-Layer Perceptron (MLP) neural network will be applied.

In Figure 4.2 a selection of the parameters that are used to prepare the MLP response parameter ζ can be seen. These include the mean scaled reduced width (MSCRW) and length (MSCRL), along with the spread of each ($\sigma_{\mathrm{MSCRW}}, \sigma_{\mathrm{MSCRL}}$), the reconstructed shower depth maximum (Xmax) and its spread (σ_{Xmax}), the spread (σ_E/E) and chi-squared fit ($\chi^2_E/(ntel-1)$) of the reconstructed energy (where the energy is derived from the total image size) and finally the time-slope across the camera (the gradient in the timing information from each pixel resulting from the light travel time across the camera). These parameters are fed into the MLP algorithm which, by use of a structure of parallel neurons, most easily compared to the function of biological neurons, learns the similarities and correlations between the

shower parameters. Thus the algorithm is able to combine this information into a single identifying classification factor for each image. An example of the resulting MLP response parameter (ζ) distribution as a function of energy can be seen in Figure 4.3. The actual cut value is derived for each energy bin and is based on maximising the difference between the signal and background significance. Following this background rejection technique, further cuts and analysis are performed in order to obtain the final performance indicators; these will be discussed further in Section 4.8 after first looking at how GCT is created within the simulation framework.

4.6 Building GCT

In initial large scale simulation productions for CTA, the model for GCT was a preliminary best estimate based on knowledge available at the time with a "MaPM like camera". In 2015, Prod3 was carried out with the specific intention of arriving at an idea of the optimal layout for CTA following the site selection. For this work there was a desire to have a more accurate depiction of GCT, specifically including a camera using SiPMs, and to this end an updated version of the SIM_TELARRAY configuration files has been created and will be presented here. A more detailed definition of the individual parameters can be found in [33].

4.6.1 Optics

The optics of a dual mirror system have been described in Chapter 3. In simulations the system is reduced to its vital components: the shadowing from the secondary mirror, masts and camera, and the shape and optical quality of the mirrors.

Mirrors

The Schwarzschild-Couder design of GCT consists of two aspherical mirrors, the shape of which has been optimised in order to achieve a good PSF on axis and well within the pixel size out to the edge of the field of view. The shape can be described by a 16^{th} order polynomial [126] which can be expressed as,

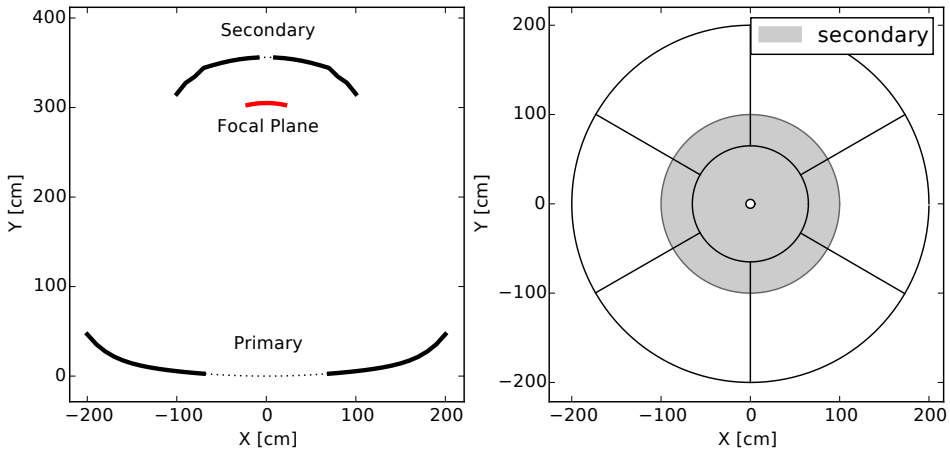


Figure 4.4: The mirror positions for GCT as viewed from the side (left panel) showing the shape of the secondary (black solid), primary (black solid) and focal plane (red solid) along with the holes in the mirrors (dashed line). The right panel represents a top down view of the mirrors highlighting the segmentation of the primary into 6 petals.

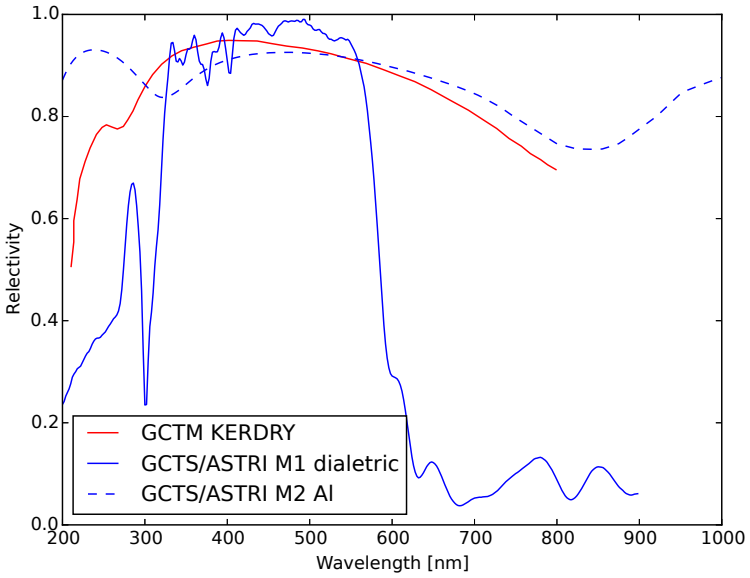


Figure 4.5: Mirror reflectivity as a function of wavelength for GCTM (red) and GCTS (blue). GCTM will use an $AlSiO_2$ coating for both mirrors while GCTS is expected to have a dielectric coating on at least the primary to reduce the effect of NSB.

	Primary	Secondary
p_2	$5.42 \cdot 10^{-4}$ m^{-1}	$2.44 \cdot 10^{-4}$ m^{-1}
p_4	$3.39 \cdot 10^{-10}$ m^{-3}	$3.88 \cdot 10^{-8}$ m^{-3}
p_6	$-1.35 \cdot 10^{-13}$ m^{-5}	$-1.31 \cdot 10^{-11}$ m^{-5}
p_8	$1.29 \cdot 10^{-17}$ m^{-7}	$2.88 \cdot 10^{-15}$ m^{-7}
p_{10}	$-6.85 \cdot 10^{-22}$ m^{-9}	$-3.98 \cdot 10^{-19}$ m^{-9}
p_{12}	$2.01 \cdot 10^{-26}$ m^{-11}	$3.34 \cdot 10^{-23}$ m^{-11}
p_{14}	$-3.06 \cdot 10^{-31}$ m^{-13}	$-1.54 \cdot 10^{-27}$ m^{-13}
p_{16}	$1.89 \cdot 10^{-36}$ m^{-15}	$2.99 \cdot 10^{-32}$ m^{-15}

Table 4.1: Pre-factors for the polynomial describing the primary and the secondary mirrors.

$$F(x) = \sum_i p_i x^i, \tag{4.6.1}$$

where x_i is the radial distance from the centre of the telescope axis and p_i is the polynomial pre-factors described in Table 4.1. This is shown in Figure 4.4 where it can also be seen that both the primary and secondary reflectors have holes in the centre and that the primary is made up of 6 petals. The mirrors are separated by a distance of 3.56 m and the telescope focal length is 2.28 m.

For each of the mirrors there is an associated reflectivity as a function of wavelength. Currently for the two camera types there are two different mirror coatings. For GCTM, using MaPMs, an aluminium quartz (Al SiO_2) coating is envisaged whereas the SiPM camera, GCTS, is expected to utilise a form of dielectric coating on at least the primary with the secondary using a Al SiO_2 coating[2]. The reason for the difference is due to the increased sensitivity of silicon devices to longer wavelengths and therefore in the night sky background region (see Figure 4.6 for a comparison of the Cherenkov spectrum to the NSB and Section 4.6.4 for further information). Therefore, in order to suppress the NSB, a spectral cut off would be

[2]Although the primary reflector is on a whole larger, it is made up of smaller petals. Therefore it is easier to coat than the monolithic secondary due to the size of the coating chambers for the dielectric method.

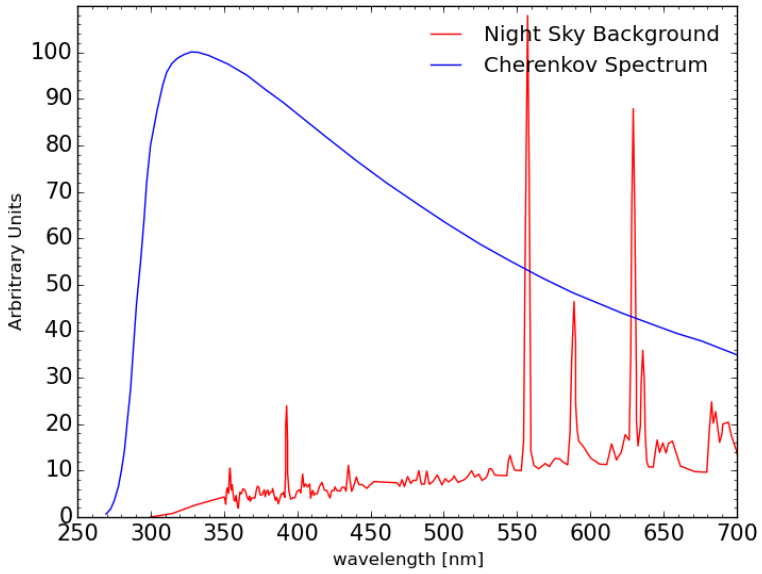


Figure 4.6: Figure showing the shape of the Cherenkov spectrum from gamma-rays compared to the measured shape of the NSB spectrum. As can be seen, above ∼550 nm there are several large peaks in the NSB which would be desirable to avoid. The two spectra have been scaled separately for easy comparison and are not representative of the absolute levels.

needed somewhere above 550 nm.

In Figure 4.5 the measured reflectivities are shown. The GCTM data were obtained from measurements performed by Kerdry Industry in France. For GCTS, due to similar design features with the ASTRI project along with the idea that the two groups are working towards the same goal, the measurements from the ASTRI mirrors have been adopted.

For each configuration there is also included a mirror reflectance random angle. This introduces a random Gaussian scatter (σ =0.0075 degrees) of the Cherenkov photons to account for any small scale surface deviations. Unfortunately all other mirror errors (such as alignment) are currently unavailable for the two mirror design and will hopefully be incorporated in the near future. For now they are treated as negligible secondary effects.

Shadowing

The Cherenkov light that is observed by the camera must first pass through the optical structure of the telescope. For a two mirror telescope this includes the camera, the secondary reflector and the masts and trusses that hold the camera and mirrors in place. The amount of light that survives makes up the total transmission with the rest defined as shadowing. The secondary mirror, which is 2m in diameter, is responsible for the majority of the shadowing (around 25% for on axis light). In SIM_TELARRAY this is taken into account separately by defining the diameter of the reflector, and the same is true for the camera. It is assumed that the camera, including the cooling unit and wind shield, can be contained by a cylinder 42 cm in diameter with a depth of 50 cm.

The remaining shadowing factor from the masts and trusses must be specified as a function of field angle out to the field of view of the camera[3]. This has been evaluated in [25] using the ray tracing program ROBAST [97] in which the light path is traced through a 3D model of the telescope; see Figure 4.7 for an example using a H.E.S.S. telescope and Figure 4.8 for the resulting reduction in collection area from different components of the structure. The total transmission passing just the trusses and masts is then modelled with a polynomial and produces a result for GCT as

$$T(\theta) = 0.881 \cdot (1 + 3.271 \cdot \sin^{1.66}(\theta))^{-1} \tag{4.6.2}$$

where $T(\theta)$ is the transmission as a function of the field angle θ.

4.6.2 Detectors

As has been discussed throughout previous sections, the major difference between GCTM and GCTS is the choice of detectors in the camera, SiPMs and MaPMs. The physical difference in these detectors results in much better single p.e. response for the SiPM, which is easily able to distinguish individual events, but the MaPMs are

[3]While it is possible for SIM_TELARRAY to perform ray tracing to include shadowing from the masts,the added computation far outweighs any observed benefits

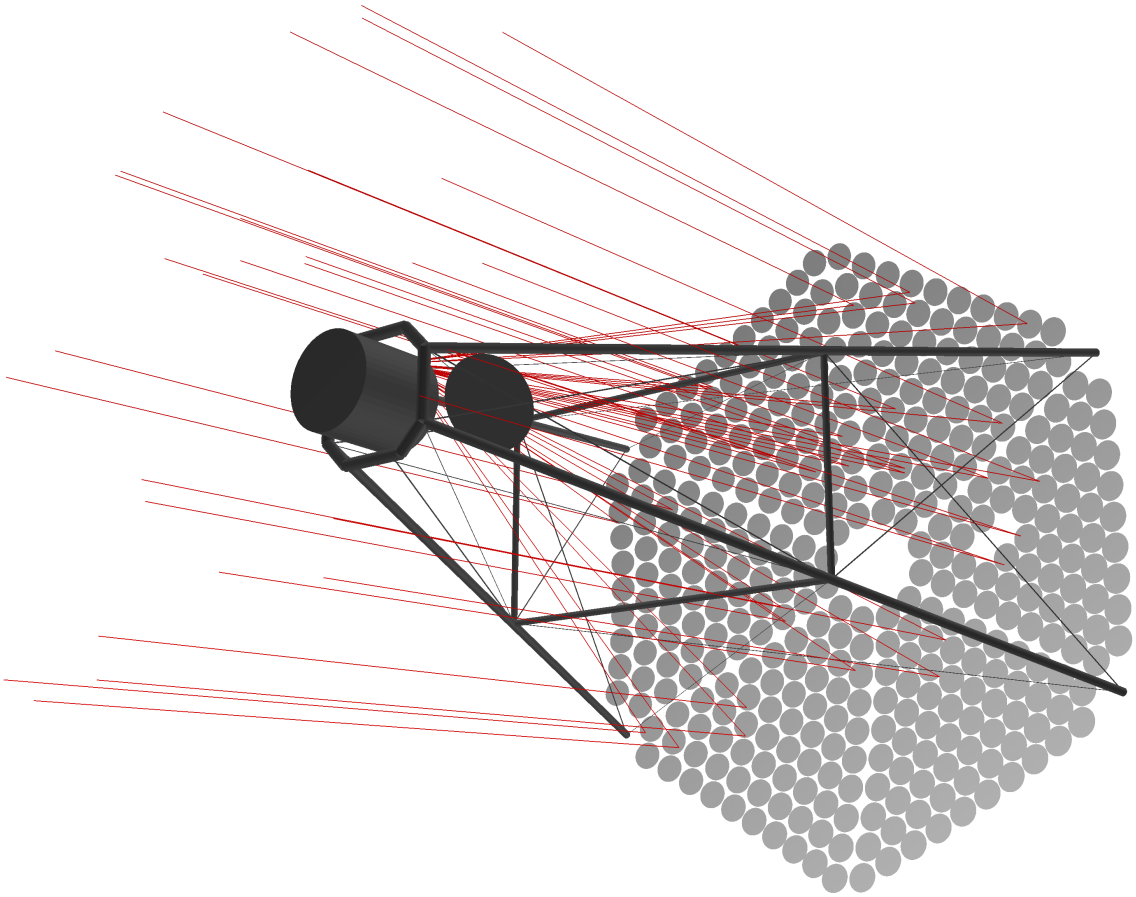


Figure 4.7: Example of ray-tracing for a H.E.S.S. type I telescope where the red lines represent the trace of the photons on their path to the camera. Image from [98].

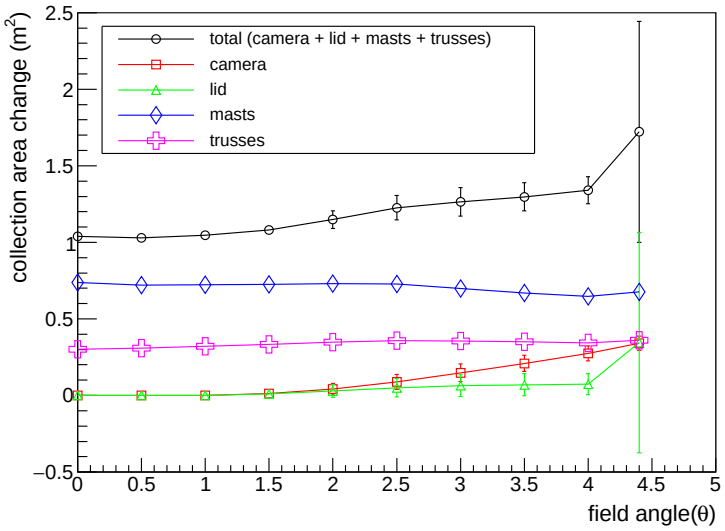


Figure 4.8: The shadowing for the separate components of the GCT structure as a function of the field angle (excluding the secondary mirror) along with the total of the camera, lid, masts and trusses. For the shadowing component needed by SIM_TELARRAY only the masts and the trusses are needed. Figure from [106].

sensitive in the optimal range for detecting Cherenkov emission without the need for spectral cut offs. Let us now consider the crucial parameters for the detectors within SIM_TELARRAY.

Focal Plane

Each camera will consist of 32 modules with a total of 2048 pixels. Each module will follow the focal plane radius of curvature, of 1.0 m. For the two configurations there are different pixel sizes, 0.6125 cm and 0.623 cm for GCTM and GCTS respectively, providing a pixel FOV of $\sim 0.15°$. As both cameras are roughly the same size, the larger pixel size of the silicon device results in smaller gaps between modules which will benefit performance. The pixels are defined to follow the curvature of the focal plane[4].

[4]Future revisions of SIM_TELARRAY may allow for pixel groupings to align with the surface of individual sensor modules.

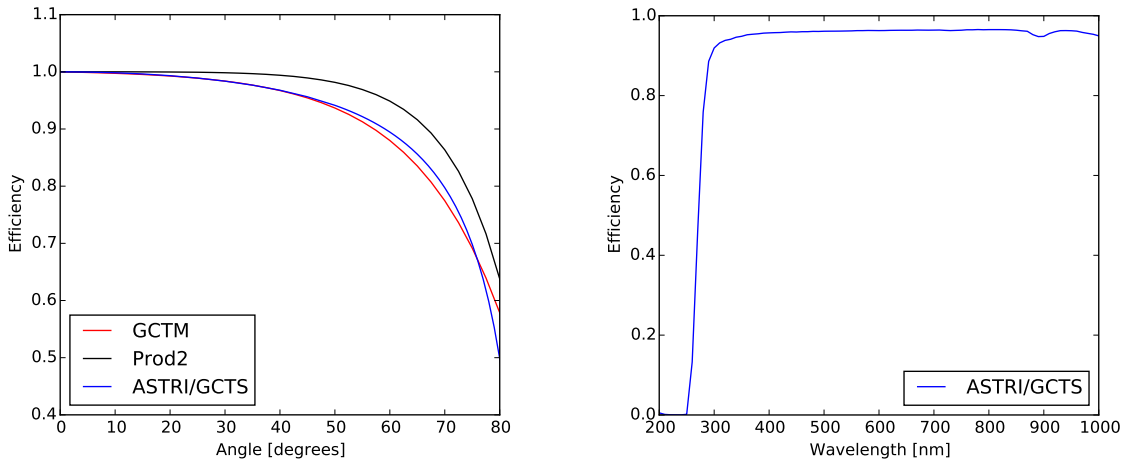


Figure 4.9: Left: The angular acceptance of the protective glass layer in the case of MaPMs (red) and the window and resin for the SiPM (blue). Also shown is the value used in past productions (Prod2,black). Right: the wavelength acceptance of the GCTS protective window for on axis events.

Windows

The active components of each detector require some protection. In the case of the MaPMs there is a glass layer that covers the front of each module. For the SiPMs there is only a layer of resin and therefore it is expected that GCTS will have a protective window over the focal plane. Both the glass layer over the MaPMs and the window for GCTS will affect the light that reaches the detector, mainly as a function of incidence angle, where photons with a large inclination have a greater probability of being reflected. In the left panel of Figure 4.9 the expected efficiency as a function of field angle can be seen, where the measurements for GCTM where obtained from [39]. For GCTS, measurements were provided by ASTRI, along with a wavelength acceptance (Figure 4.9 right) for a Plexiglass G-UVT acrylic sheet window.

Photon Detection Efficiency

The two detector types convert photons into p.e. via different processes and therefore have distinctive efficiencies for this. The Cherenkov spectrum starts to peak from around 290 nm and then slowly drops off, while the NSB spectrum starts becoming

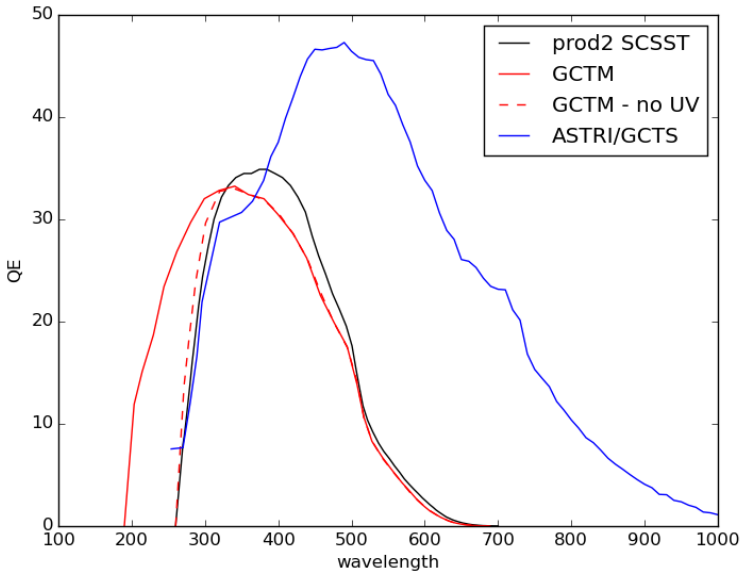


Figure 4.10: Photon detection efficiency for GCTM (red) with and without an enhanced UV filter and GCTS (blue). Also shown is the data used in previous productions (Prod2, black).

a problem past 550 nm (recall Figure 4.6) and it is therefore over the ∼300-550 nm range that detection efficiency should be maximised. In Figure 4.10 the photon detection efficiency (PDE) curves for different configurations are shown. For GCTM there are two options with and without an improved UV response achieved by way of a UV enhanced glass as opposed to the standard Borosilicate glass. The new device chosen for GCTM is the Hamamatsu H12700A and is shown along with the H10966A, which was previously used in older simulations (and is currently in the camera prototype). Although the absolute PDE is slightly lower for the newer device, it boasts an improved single p.e. response (see following section). The data are taken from the published data sheets for each device.

For GCTS, as with the mirror reflectivity, the data were provided by ASTRI and are based on measurement for the Hamamatsu LCT5 (70 μm cell pitch) device.

It is also important to take into account fluctuations in the PDE. For a realistic detector, it is expected that the PDE (4% for each camera) and gain (20% for GCTM and 5% for GCTS) will vary for each module. GCTM is also expected to experience a voltage variation of 3%.

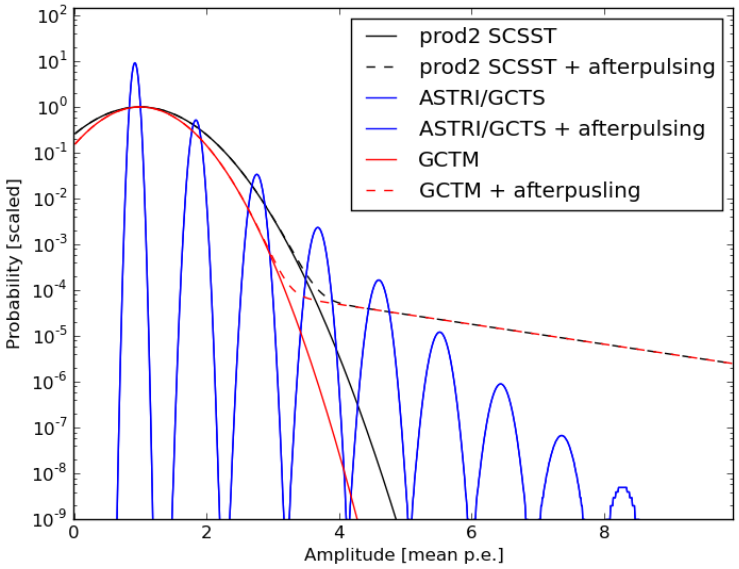


Figure 4.11: Single photo-electron (p.e.) response for GCTM (red) and GCTS (blue). The new device chosen for GCTM has an improved response over that used in previous productions (black). The dashed lines showing the after pulsing which is an effect of delayed ions interacting with photo-cathodes arriving within the readout window and is only applied to p.e. originating from the NSB.

Single Photo-Electron Response

If PDE can be considered as the efficiency of the device for the conversion of photons into an electronic signal, then the single photo-electron (SPE) response indicates how accurately the device does so. In a traditional PMT in which a photon produces an electron within the photocathode, this is then accelerated down the tube, producing secondary electrons at each cathode. The end result is an amplified signal which can be measured. However, the output signal will vary due to the probabilistic nature of the secondary electron production process. This will generally produce a Poissonian distribution around the expected value. The same concept holds true for MaPMs and the expected distribution can be seen in Figure 4.11. Also shown is the much more distinctive SPE response for the silicon device, which relies on the production of electron-hole pairs. The resulting charge created is then read out and amplified, a process that produces a more reliable estimate of the true charge. In Figure 4.11, for GCTM there is also an additional component referred to as the

afterpulse. This is an effect caused by the registering of NSB-produced ions within the photocathodes, which, as the NSB represents a constant noise, will appear within the readout window effectively creating a non negligible tail (the level of the NSB will be calculated in Section 4.6.4). A signal produced by Cherenkov photons would produce a similar effect; however, this usually occurs outside the readout window, therefore the data without after pulsing are sampled for Cherenkov signals while NSB induced signals sample from the data with after pulsing.

The data for the new GCTM detectors were obtained from measurements provided by members of the CLAS12-RICH project (a ring imaging Cherenkov detector) [51][private comm.], and represents an improved response compared to the current H10966A devices. The SiPM data again were provided by the ASTRI group. For GCTS the after pulsing effect is not included; future work is planned to correctly model the SiPM SPE to account for after pulsing and optical cross talk (noise between pixels).

Finally, for each p.e., the pulse amplitude is sampled randomly from the SPE response and shifted acording to the photon arival time, the transit time through the detector (5.3 ns for GCTM and 4 ns for GCTS) and a random level of noise resulting from jitter within the detector (0.28 ns for GCTM and 0.2 ns for GCTS).

4.6.3 Electronics

The digitisation of the signal, different trigger levels and the readout all fall under the heading of electronics as far as the simulations are concerned. In the camera, the electronics consist of the Target modules, backplane and DACQ boards (See section 3.5.2).

Trigger and Discriminator

Once the shower p.e. have been obtained for the camera, they need to be converted to an electronic signal in mV. This, along with the first level trigger, is performed

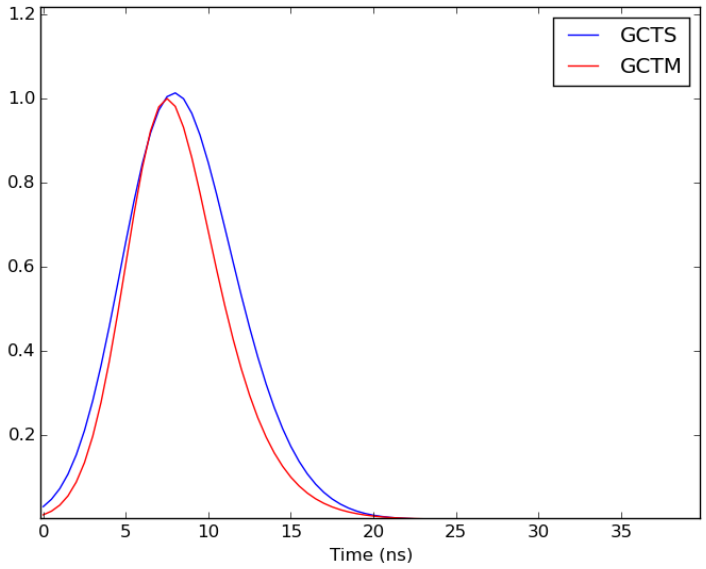


Figure 4.12: Normalised discriminator pulse shape for GCTM (red) and GCTS (blue). This is used to shape the input signal before being passed to the discriminator logic.

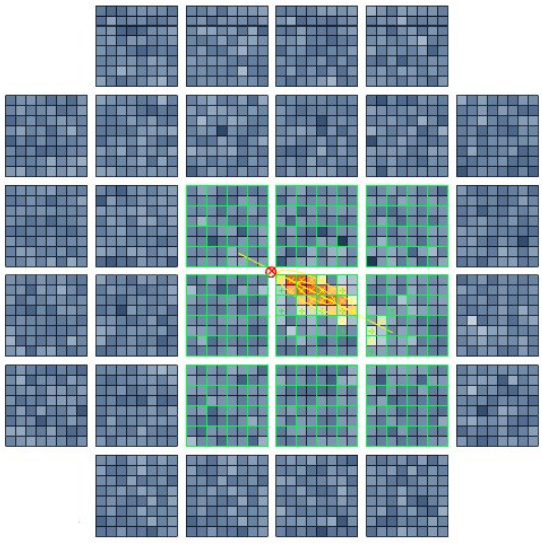


Figure 4.13: Focal plane showing a triggered shower image. The green squares represent a grouping of four individual pixels to form a super-pixel to which the discriminator logic is applied.

in the simulations by way of a discriminator[5]. The pulse shape at the input of the discriminated can be seen in Figure 4.12 and is assigned an amplitude, which for both GCTM and GCTS is defined to be 20 mV (signal amplitude after amplification per mean p.e.). The pulse shape for GCTS is slightly wider (∼20%) than that for GCTM due to the pulse shaping performed by the silicon pre-amplifier.

This amplified signal is then passed to the discriminator logic, requiring the signal within a pixel, or set of pixels, to be above a certain threshold for a given amount of time and/or integrated signal. The determination of the thresholds will be considered in Section 4.6.5. The output from the discriminator consists of a 42 mV pulse with a rise and fall time of 1 ns.

For GCT, the trigger logic requires 4 neighbouring pixels, a so-called "super pixel", to be above a given threshold, an example of which can be seen in Figure 4.13. This is a hard coded majority trigger and must be specified explicitly in the configuration file based on pixel numbers. In order to flag a trigger, at least two neighbouring super pixels must meet the threshold.

Digitisation and Readout

Once the camera has triggered, the data must be digitised and read out. For the simulations this is represented by the process of a Flash Analogue to Digital Converter (FADC). The current hardware design for GCT allows for a sampling speed of up to 1 GHz, which can be reduced depending on the desired data rate. The pulse shape at the input of the FADC is the same as that of the discriminator (Fig 4.12) and the signal is then divided into 128 time intervals, of which 96 will be summed up starting 24 bins before the trigger time. The resulting amplitude in each time slice for a single p.e. is 8 ADC counts with a pedestal of 40. Gaussian noise is also added to each bin, giving a signal to noise ratio of 8:2 ADC counts.

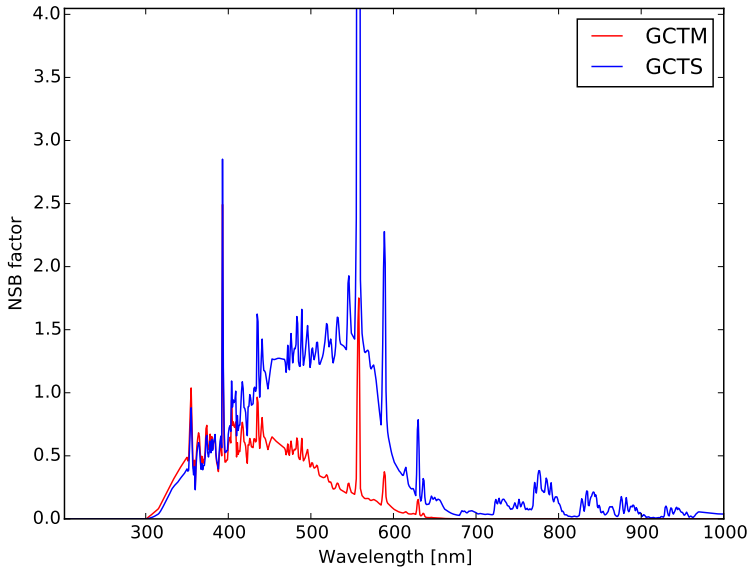


Figure 4.14: Convolved night sky background spectrum accounting for the individual efficiency parameters within the telescope configuration applied to measurements taken from La Palma.

4.6.4 Night Sky Background

The expected background from the NSB is derived from the spectral shape of measurements taken in La Palma [30] scaled with the recorded value observed by H.E.S.S. in Namibia[6]. In SIM_TELARRAY, the NSB is given as a single value which is then added as random white noise. The NSB therefore needs to be converted from a measured flux to a rate per pixel. This is done with the TESTEFF script that is available with the CORSIKA_SIMTELARRAY package. This takes into account the configuration of the telescope and atmospheric transmission and convolves these with the NSB spectrum, the result of which can be seen in Figure 4.14. Integrating this provides a NSB light factor which is then converted into a scaled value taking account of the telescope optics

[5]The use of a discriminator in the sim_telarray terminology is a relic from the days of the HEGRA experiment.

[6]Although there will be differences in the NSB at different CTA sites, exact values are not available and therefore the H.E.S.S. scaled value is assumed for all possible observatory locations. This is reasonable for a southern site that is exposed to the Galactic plane (produces a large amount of NSB).

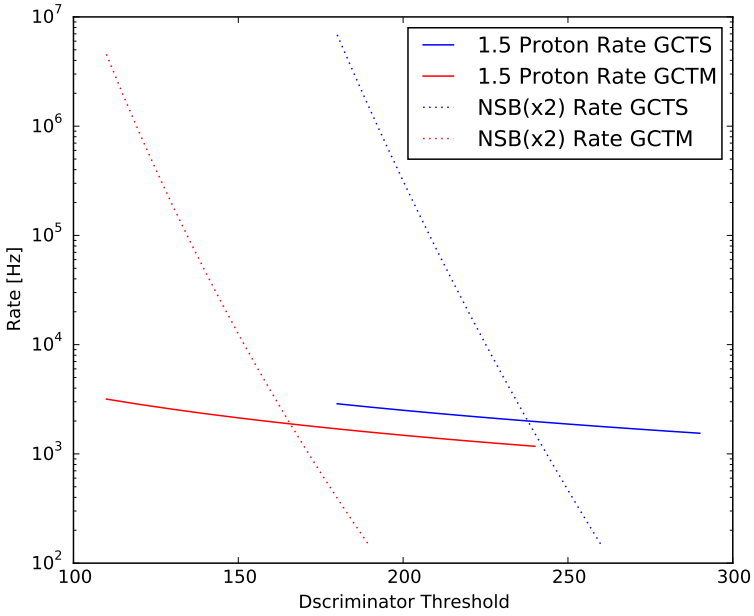


Figure 4.15: Determination of a safe discriminator threshold, obtained by analysing runs with an increasing discriminator threshold and measuring the trigger rate. For the protons the data are obtained from a number of CORSIKA proton showers whereas the NSB rate is determined by running SIM_TELARRAY with a dummy input containing no showers. The safe operational threshold is defined by eqn 4.6.5 and equates to where these lines intersect.

$$N = \frac{A \cdot P}{f^2} \cdot \int_{\lambda_1}^{\lambda_2} S(\lambda) d\lambda, \tag{4.6.3}$$

where A is the mirror effective area, P is the pixel area, f is the focal length and $S(\lambda)$ is the convolved NSB spectrum (appropriate values of λ_1 and λ_2 should be chosen in order to cover the non zero range of the convolved NSB spectrum, i.e. 300 - 1000 nm). Finally, in order to convert into an expected NSB rate, a correction factor is derived from the known NSB rate for H.E.S.S.

$$NSB = NSB_{\text{HESS}} \cdot \frac{N_{GCT}}{N_{HESS}}, \tag{4.6.4}$$

which equates to 14.2 MHz and 41.0 MHz for GCTM and GCTS respectively.

4.6.5 Discriminator Threshold

In order for the camera to trigger, as described previously, two super pixels (sum of 4 pixels) must have a value greater than a given threshold. This is required in order to obtain a sensible trigger rate that is unaffected by fluctuations of the NSB. The safe discriminator threshold is defined to be at the point where

$$\text{Rate}_{2\times\text{NSB}} = 1.5 \times \text{Rate}_{\text{proton}}, \tag{4.6.5}$$

i.e. the rate of events triggered by protons should be 1.5 times greater than the spurious triggers created by double the operational NSB. To find this value a set of simulations has been carried out using protons and dummy events with no showers over a range of discriminator thresholds. The result can be seen in Figure 4.15 where the NSB rate drastically decreases with an increasing discriminator threshold compared to a moderate decrease in the proton rate. The point where these intersect is set as the safe threshold. This equates to a level of 172 mV for GCTM and 230 mV for GCTS.

4.6.6 Summary

Only a general overview of the most important configuration parameters has been presented. A summary diagram can be seen in Figure 4.16 and a full table of all relevant parameters can be found in Appendix B. The total chain can be considered as the following. From the Cherenkov light that intersects with the fiducial CORSIKA sphere, the amount remaining is governed by the set of efficiencies,

$$\epsilon(\lambda) = S \cdot R(\lambda) \cdot R'(\lambda) \cdot E(\lambda) \cdot Q(\lambda) \tag{4.6.6}$$

where S is the shadowing factor, $R(\lambda)$ and $R'(\lambda)$ are the reflectivities of the primary and secondary reflectors, $E(\lambda)$ is the efficiency of the camera window and $Q(\lambda)$ is the photon detection efficiency. Once this has been obtained, a total signal is built up from the individual photons, sampling from the measured SPE, giving

$$\text{Signal} = \sum SPE_{\gamma} + \sum SPE_{\text{NSB}} \tag{4.6.7}$$

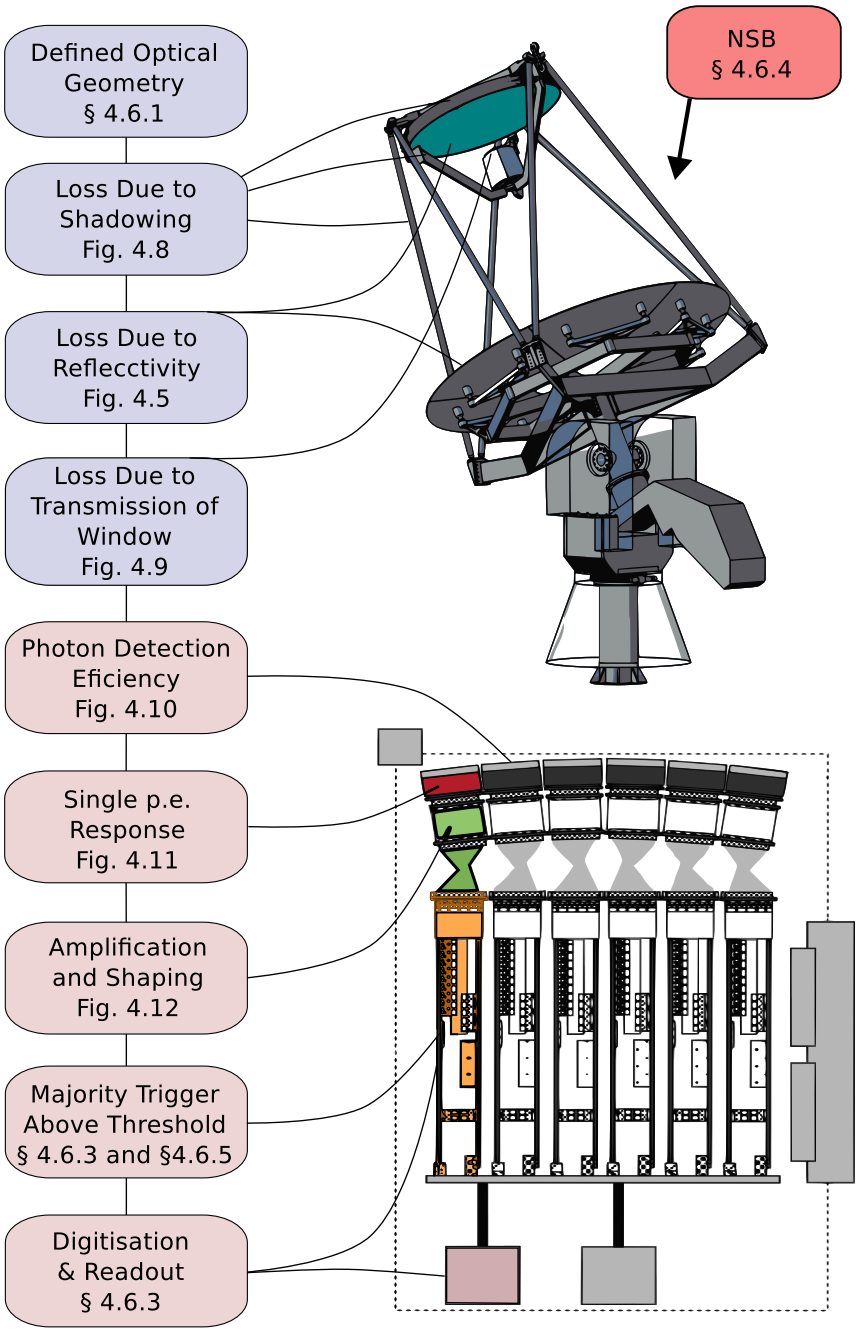


Figure 4.16: Summary of the important configuration parameters with the top 4 related to the structure and the bottom 5 to the camera. A more detailed diagram showing the structure and camera can be seen in Figures 3.13 and 3.15.

which will then be amplified and shaped according to the discriminator pulse shape at the start of the discriminator logic. The first level trigger then consists of a requirement that the sum of 4 neighbouring pixels be above the discriminator threshold. A second level trigger requires 2 or more super pixels to exist above the threshold and finally a third level trigger may be applied to telescope multiplicity in order to obtain stereo reconstructed events. Having passed all these requirements the data are then read out as a GHz sampled 96 ns long waveform for each pixel.

4.7 Low Level Evaluation

Having created an updated configuration for GCT, it is important that the output be correctly evaluated. In this section the first, or "low level" checks that can be performed in order to check the performance of the telescope will be shown. For the majority of these only the response of an individual telescope, or, where a stereo trigger needs to be included, two telescopes must be considered.

4.7.1 Image Amplitude Trigger Efficiency

The amount of observed light, and therefore the number of p.e., within the camera required to form a trigger is an important indicator of efficiency. It should expected that, for a certain minimum number of p.e., a reliable trigger should be obtained. This in fact forms a requirement within CTA[7] which states that a trigger probability of at least 50% must be achieved with an image amplitude of at least 100 p.e. and a goal of less than 80 p.e.. Therefore both GCTM and GCTS are considered and the trigger efficiency to the total p.e. within the camera is compared[8]. The result can be seen in Figure 4.17 where a trigger efficiency of 50% at 60 p.e. for GCTM and 72 p.e. for GCTS is achieved. Both configurations meet the requirement and even achieve the goal.

[7]For internal CTA reference: B-SST-1230

[8]The origin of the light is not an important factor, but in this case we use primary gamma-ray showers

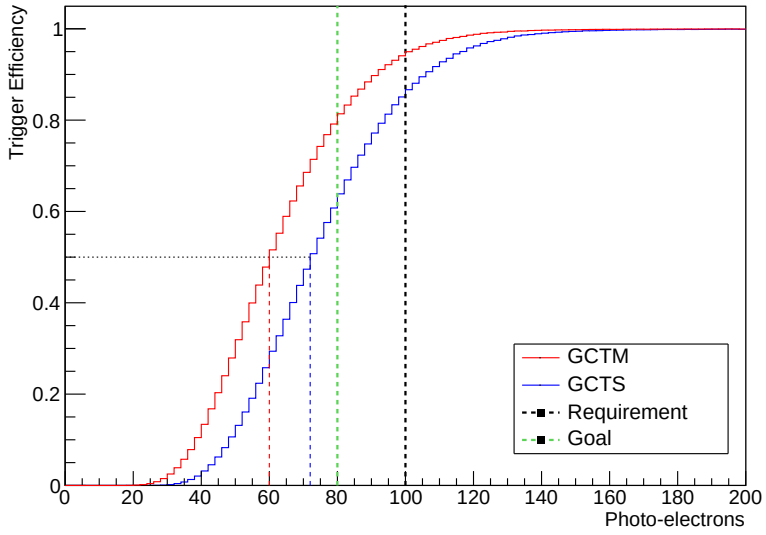


Figure 4.17: Trigger efficiency as a function of image amplitude (p.e.). The black and green dashed vertical lines represent the required and goal amplitude where a 50% trigger efficiency is obtained. Both GCTM (red) and GCTS (blue) satisfy the goal by obtaining a 50% trigger efficiency at 60 p.e. and 72 p.e. respectively.

4.7.2 Charge Resolution

The conversion of Cherenkov light into measured intensity by the camera is possibly the most important task for the system. The total charge reconstructed by the camera in each pixel, Q_{rec}, should therefore be comparable to the simulated amount of Cherenkov light, Q_{true}, expressed in p.e.. In order to evaluate this, the fractional charge resolution is defined to measure the spread of the reconstructed charge around the simulated charge such that

$$\frac{\sigma_Q}{Q} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}(Q_{rec,i} - Q_{true})^2 - Q_{true}}}{Q_{true}}. \tag{4.7.8}$$

The benefit of expressing the resolution in such a form is that it simultaneously takes into account the spread of the reconstructed charge and also any possible bias that is present.

For the camera, it is possible for the whole waveform for each event to be recorded, equating to a 96 ns trace for each pixel. However, the expected signal from a shower would be much shorter than this, therefore it is necessary to integrate

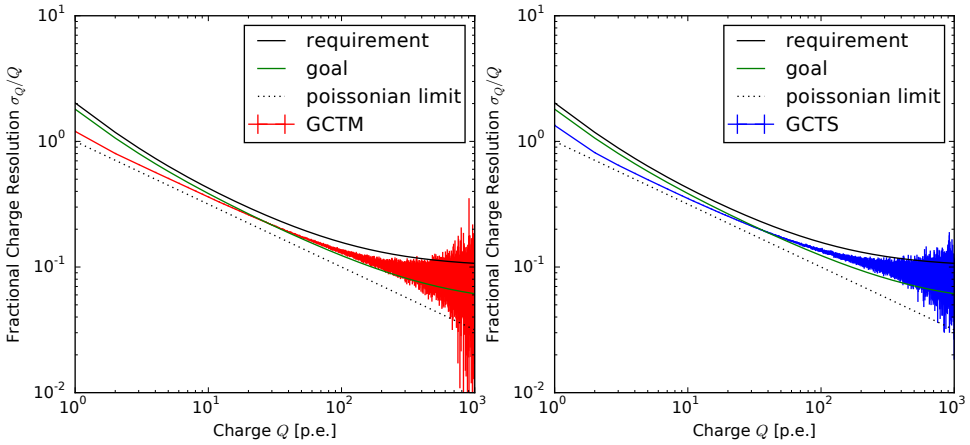


Figure 4.18: The charge resolution for the chosen integration scheme (local peak) for GCTM (left) and GCTS (right). Also shown is the required and goal charge resolution as defined by CTA along with the Poissonian limit which expresses the best obtainable charge resolution.

this trace over a shorter window around the region of interest. This will allow a reliable estimate of the total charge to be determined, removing effects from electronic noise and NSB.

A very simplistic method to select a region of interest, known as *local peak* integration, would be to search for the maximum signal within an individual trace and integrate about that point. Alternatively, the choice of integration point can be taken from an average of all peak positions of an event, known as *global peak* integration, or from a subset of adjacent pixels, the *nearest neighbour* integration. For high energy showers observed at large distances, there will be a large time spread across the camera. Therefore methods such as *global peak* would perform poorly. In Figure 4.18 the charge resolution for both GCTM and GCTS, using the local peak method, can be seen. Also shown are the required and goal charge resolutions for CTA[9]. Both cameras meet the requirement which can be seen as the black dashed line and is based on an RMS uncertainty from a background of 0.125 p.e./ns.

[9]For internal CTA reference: B-SST-1010

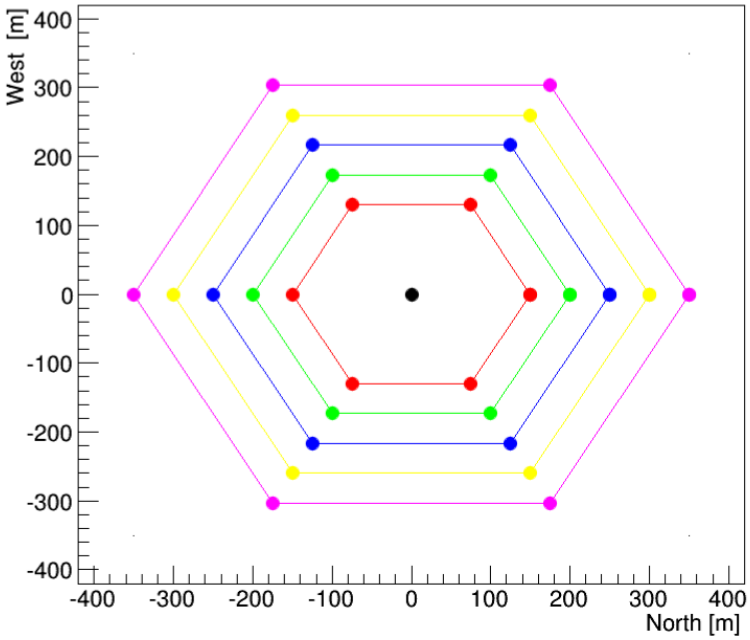


Figure 4.19: MC SST mini production of 5 arrays of 7 telescopes with increasing inter-telescope distance from 150 m to 350 m.

4.8 High Level Evaluation: Mini Array

To further understand the potential of GCT, it is important to consider the performance as a small array of telescopes. By doing this it is possible to gain an insight into the potential science that GCT could achieve as part of an mini array of telescopes operating early in the construction phase of CTA.

4.8.1 MC SST Mini-Production

The current forecast for CTA is that the Small Size Telescope (SST) component of the southern observatory in Chile will comprise roughly 70 telescopes, most likely made up of a combination of GCT, ASTRI and SST-1M. While the large scale productions (Prod. 1,2 and 3) contain a variety of different layouts, the Monte Carlo SST group created a database of CORSIKA showers for a set of 5 small arrays each with 7 telescopes. This data set allowed for more in depth comparisons and collaboration within the SST group. The arrays are each arranged in a hexagonal layout differing by the inter-telescope separation which ranges from 150 m to 350 m.

	gamma-ray	protons
E_{min}	315 GeV	315 GeV
E_{max}	330 TeV	600 TeV
Γ	-2	-2
Zenith angle	20°	20°
Direction	North	North
View Cone	0°	10°
Altitude	1.650 km	1.650 km
Radius on ground	1 km	1.2 km
Re-sampling	10	20
Total Number of events	10×10^6	200×10^6

Table 4.2: Selected input parameters in the mini array production showing the energy range (E_{min} to E_{max}), the spectral index with which the events are simulated (Γ), the zenith angle and pointing direction, the viewcone from which the events were simulated (point source for gamma-rays and a diffuse source that covers the telescope FOV for protons), the altitude, the radius on the ground within which the impact points of each event are contained, the number of times each event is sampled (moved about within the radius on the ground) and the total number of events including the re-sampling.

Figure 4.19 shows the layout of the simulated arrays and in the following section, performance estimators will be derived for each, assuming on axis observations of a point like gamma-ray source and a diffuse proton background. The input parameters used to create the data set can be seen in Table 4.8.1.

4.8.2 Reconstruction and Initial Cuts

The image and shower reconstruction of events was performed with READ_CTA. This step also allows the implementation of initial cuts on the data, from basic image cuts to higher level reconstruction cuts. In this analysis, as the background rejection is performed by way of the TMVA neural network, only low level cuts have been applied. These cuts represent only initial estimates and more appropriate values will require further work to optimise the resulting performance. Therefore the cuts that were applied consist of a tail cut at a default level of 5, 10 (harder

than past analysis performed with the SSTs) and a minimum number of pixels and p.e. remaining in the image pixels post tail cuts of 5 pixels and 40 p.e..

4.8.3 Angular Resolution

The angular resolution of an instrument governs how well the structure of a source can be resolved. This is extremely important for removing potential source confusion and identifying emission regions or structure of extended sources. When observing a point source there are many factors that can lead to deviations in the reconstructed angle; these will include atmospheric effects and fluctuations in the instrument's performance (see following Chapter, Section 5.7 for more details). In practice, the angular resolution is defined as the circle in which 68% of events are contained (commonly 80% and 95% containment levels are also used but in line with CTA requirements the 68% level will be presented). This is evaluated in individual energy bins and the minimum radius required to contain the required percentage of events is returned. In Figure 4.20 the angular resolution for each of the 5 arrays of 7 telescopes is shown for all the configurations. The general trend is for an improved resolution at higher energies due to the improved reconstruction resulting from brighter showers. It can be seen that the arrays with larger spacing perform better at high energy while a more compact array will improve the low energy performance. This is largely due to the increased likelihood of observing the showers in more than two cameras when telescopes are placed close together, while the larger spacing improves the stereoscopic effect that allows better reconstruction of the shower. It can also be seen in Figure 4.20 that the angular resolution for each of the arrays is inferior to that of the predicted complete southern CTA array; this will be addressed in Section 4.9.

Once the determination of the angular resolution has been completed, the data are subjected to a further cut based on the angle between the shower and the viewing direction, θ. Traditionally the cut requires θ^2 to be less than the square of the angular resolution, due to the trigonometric shower reconstruction geometry, however in this work the cut has been drastically relaxed for two reasons. Firstly, due to the small scale of this production, the number of background events remaining past this cut

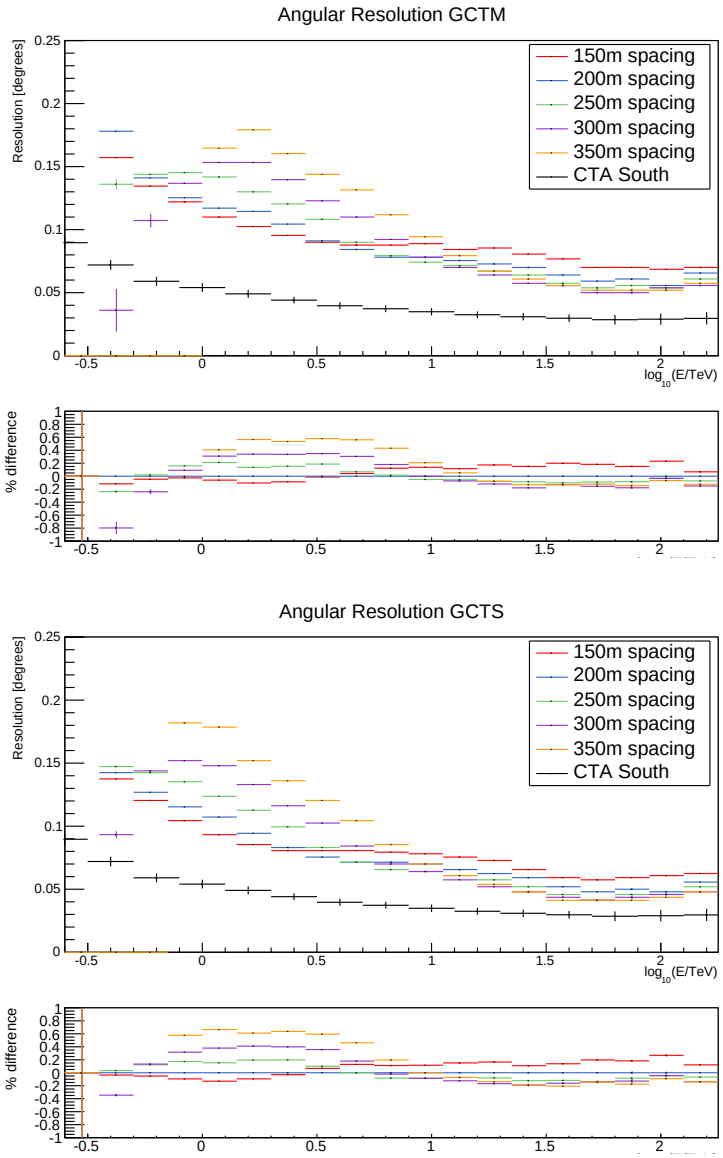


Figure 4.20: Angular resolution derived for each mini-array layout for GCTM (top) and GCTS (bottom). In each, the expected angular resolution obtained by the southern CTA observatory is also shown. The lower panels of each plot indicates the relative performance of each array compared to the 200 m spacing array (below 0: better than the 200 m array).

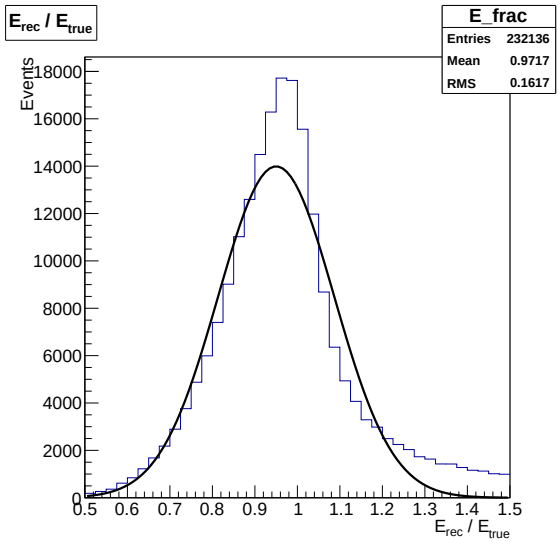


Figure 4.21: Example Gaussian fit of E_{rec}/E_{true} used to determine the energy resolution. This is performed for each energy bin where the offset also indicates the bias present in the reconstruction.

is generally too low to successfully estimate the sensitivity. Secondly, it is expected that a large number of sources that will be observed by the SSTs will be diffuse in nature, therefore a strict θ^2 cut would also remove desired events. To accommodate these two points, the θ^2 cut has been relaxed to a value of 1 square degree.

4.8.4 Energy Resolution

Akin to the angular resolution and its indication of the spatial response is the energy resolution, which reveals the spectral performance and how well the system can resolve the energy of the primary shower particle. From the reconstructed impact position and the measured lateral distribution of Cherenkov light, the energy of the incoming gamma-ray can be calculated. As with any instrument there will always be some uncertainty that arises in this measurement and this is defined as the energy resolution. In order to determine this, a comparison between the true energy and that which is calculated from the output of the system must be made. The effect can be considered as a Gaussian spread of the reconstructed energy about the true energy (E_{rec}/E_{true}, an example of which can be seen in Figure 4.21 where any offset represents a bias in the reconstruction. In Figure 4.22 the resulting energy resolution and bias are shown for each array and each configuration and it can be seen that,

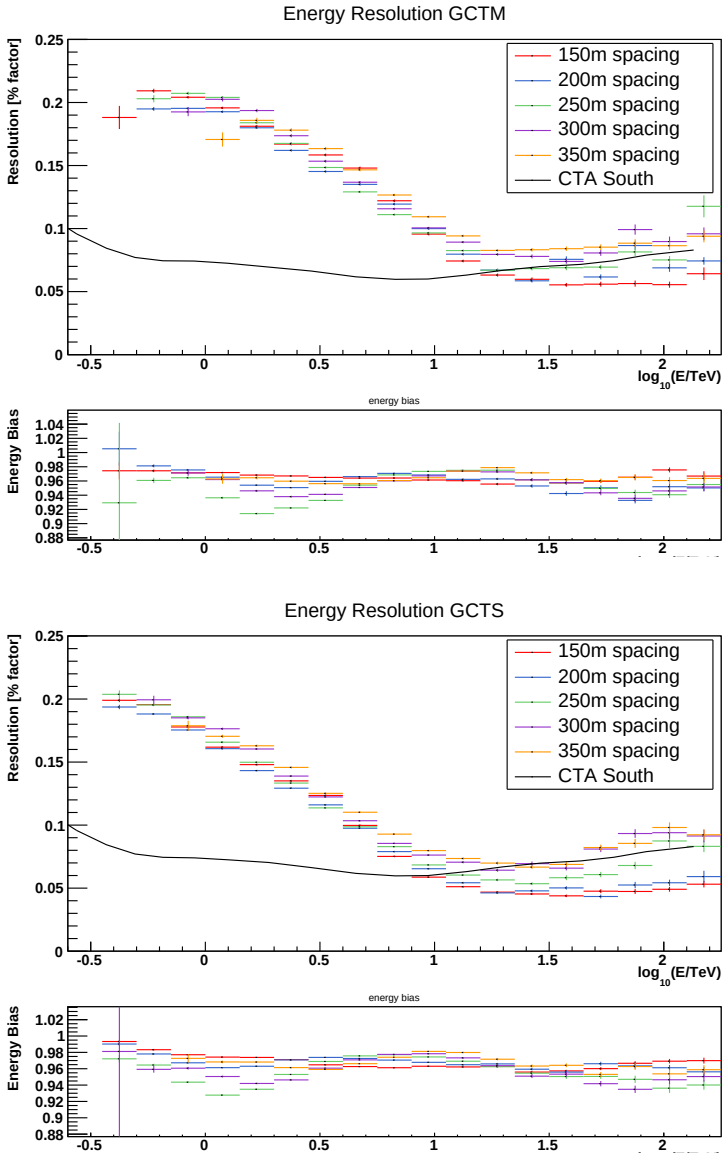


Figure 4.22: Energy resolution derived for each mini-array layout for GCTM (top) and GCTS (bottom). In each, the expected energy resolution obtained by the southern CTA observatory is also shown. The lower panels of each plot show the energy bias for each layout.

even for the mini array, each layout performance is comparable with that of the complete southern CTA array above 10 TeV. Below this energy the resolution drops off and there is an increase in the energy bias for all but the most compact arrays.

4.8.5 Effective Area

The key difference that CTA will have over previous ground-based gamma-ray instruments is the sheer amount of ground area that it will cover ($\sim$ 3 km^2 for the southern array). By folding in information about each telescope's camera and mirror efficiency, this area can be quantified by the definition of an effective area (A_{eff}). In simulations, this is calculated as the efficiency of triggering events within a given area, such that

$$A_{eff} = \frac{N_{\text{trig}}}{N_{\text{total}}} \cdot \pi \cdot R^2, \tag{4.8.9}$$

where N_{trig} is the number of triggered events passing all cuts, N_{total} is the total number of simulated events and R is the area within which shower events have been simulated. In the case of the mini array this is 1 km for γ-ray and 1.2 km for proton showers. In Figure 4.23 resulting effective area for each of the mini array layouts can be seen. It is clear that for low energy showers a smaller spacing results in an improved trigger efficiency whereas at high energies the converse is true.

4.8.6 Energy Threshold

The SST component of CTA is designed to cover the highest energy range, from around 1 TeV to well above 100 TeV. In the previous section it was shown that this is in general true, but it is important that the lower bound is determined explicitly. To do this, the concept of an energy threshold is introduced and is defined as the turnover point at which the ratio of triggered events to simulated events no longer follows the assumed observed spectrum. In Figure 4.24 the trigger fraction has been multiplied by the Crab gamma-ray spectrum with a spectral index of -2.57 for each telescope layout and the maximum gives the energy threshold. For GCTM this ranges from $\sim$ 840 GeV to $\sim$ 2.37 TeV and from $\sim$ 595 GeV to $\sim$ 1.68 TeV for

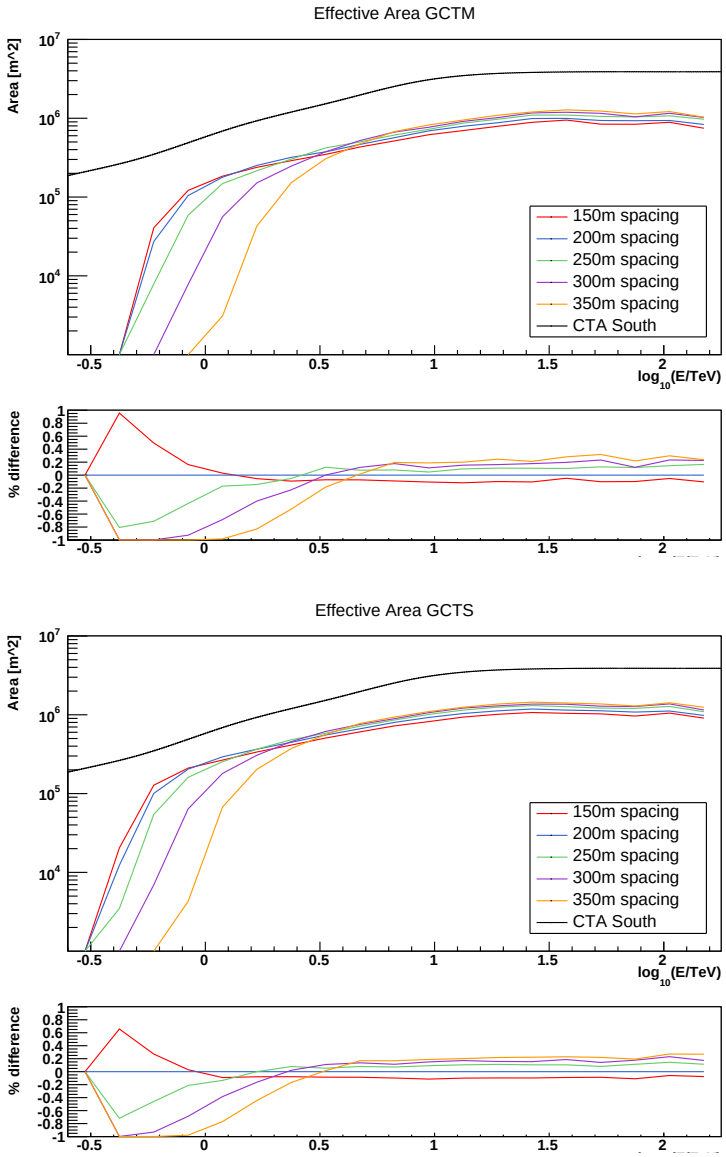


Figure 4.23: Effective area for each mini-array layout for GCTM (top) and GCTS (bottom) shown along with the expected CTA southern effective area. The lower panels show the relative difference compared to the 200 m spaced array.

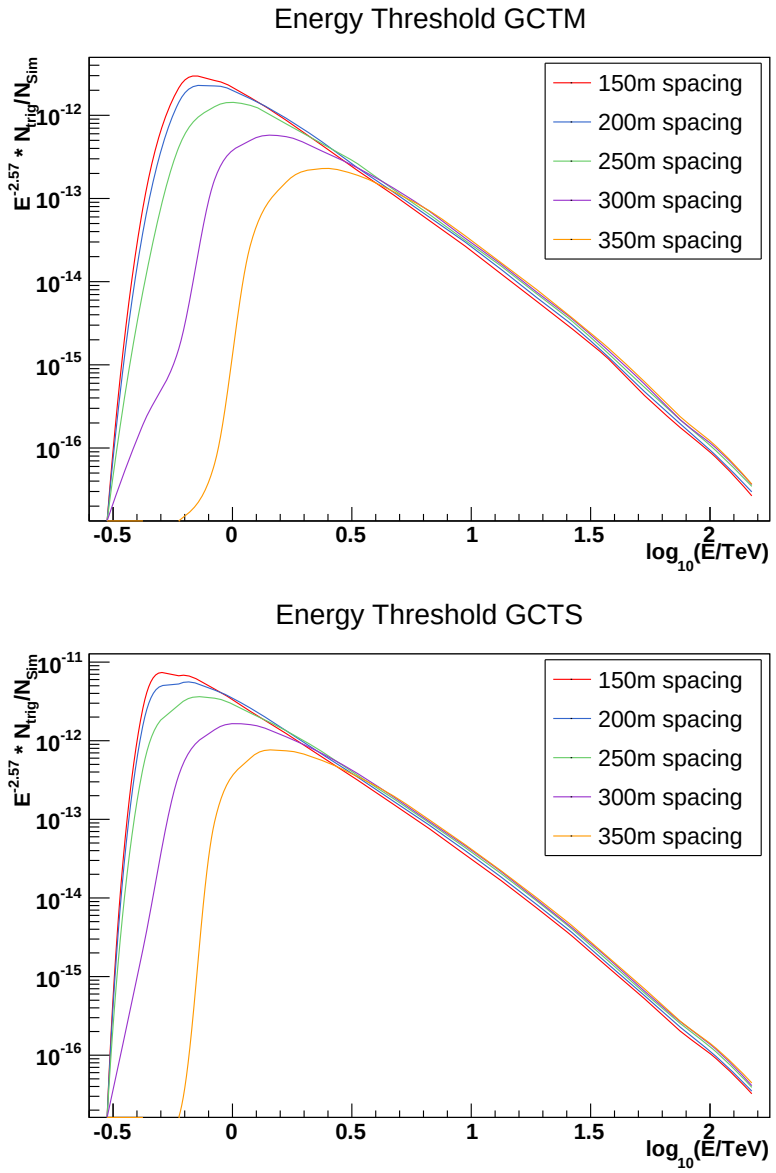


Figure 4.24: Energy threshold derived for each mini-array for GCTM (top) and GCTS (bottom). This is effectively the trigger efficiency compared the original observed spectrum, in this case assumed to be the Crab spectrum.

GCTS with increasing inter-telescope separation. Due to the larger gaps in GCTM, it is understandable that there would be a drop in performance for low energy events, which appear smaller in the camera focal plane.

4.8.7 Sensitivity

The final performance indicator is arguably the most important: the flux to which the instrument is sensitive. In general the sensitivity is defined to be the flux level at which the signal can be determined over the background to a significance of 5σ in 50 hours, based on the likelihood equations in [85]. Now, having arrived at a data set of gamma-rays originating from a point like source with a diffuse proton background remaining after the neural network rejection, the calculation needed to arrive at a final sensitivity can be made. The equation that is used to determine the significance is as follows:

$$\sigma_{\text{Li\&Ma}} = \sqrt{2N_{\text{ON}} \ln\left(\frac{1+\alpha}{\alpha} \cdot \frac{N_{\text{ON}}}{N_{\text{ON}}+N_{\text{OFF}}}\right) + 2N_{\text{OFF}} \ln\left((1+\alpha) \cdot \frac{N_{\text{OFF}}}{N_{\text{ON}}+N_{\text{OFF}}}\right)}, \tag{4.8.10}$$

where N_{ON} is the number of counts obtained in the ON region, N_{OFF} is the number of events in the OFF region and α is the normalisation between the two regions, with the OFF region defined as being five times larger than the ON region which in turn is based on the derived angular resolution (see Figure 4.26 for a representation of how these are determined). In order to calculate the number of events in relation to the observation period, the rate of detected gamma-rays and proton showers must be calculated, which in turn requires an initial spectrum to be assumed. In the case of primary gamma-rays, the spectrum for the Crab is widely assumed and takes the form

$$\phi_\gamma(E) = 2.79 \times 10^{-11} \cdot E^{-2.57} \quad \text{TeV}^{-1}\text{m}^{-2}\text{s}^{-1}\text{sr}^{-1}, \tag{4.8.11}$$

while the proton spectrum is assumed to take the measured cosmic ray spectrum simplified as a power law in the form of

$$\phi_p(E) = 9.6 \times 10^{-2} \cdot E^{-2.70} \quad \text{TeV}^{-1}\text{m}^{-2}\text{s}^{-1}\text{sr}^{-1}. \tag{4.8.12}$$

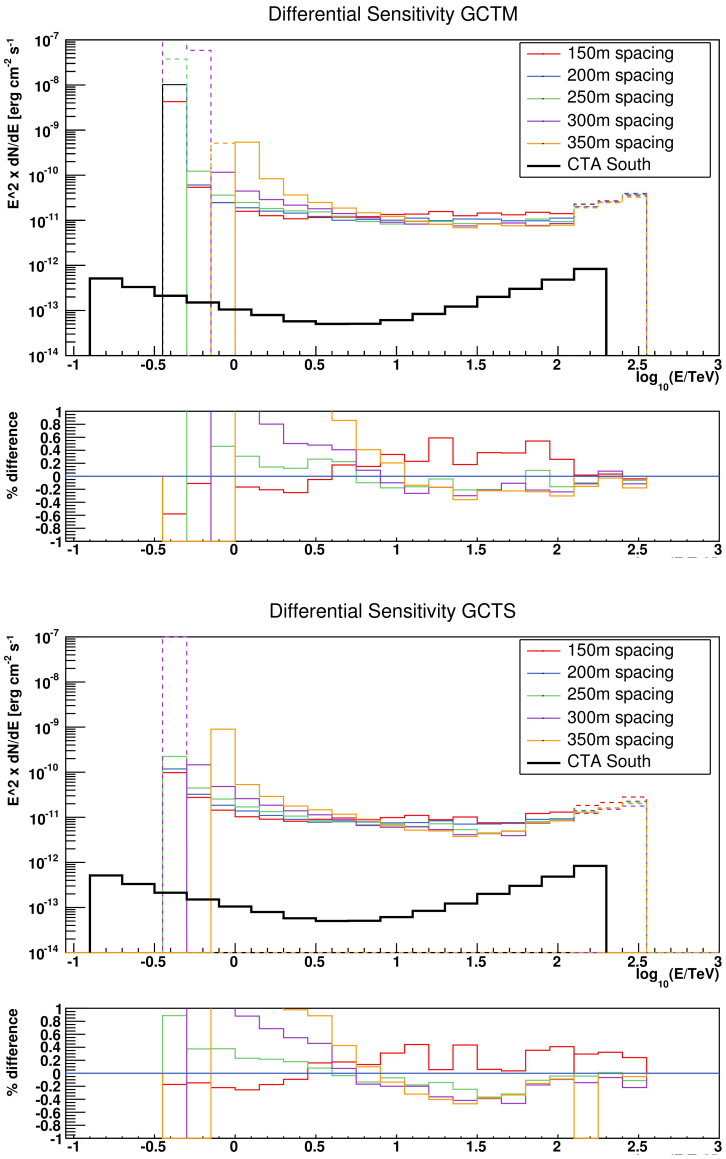


Figure 4.25: Differential sensitivity for 50 hours observation derived for each mini-array for GCTM (top) and GCTS (bottom). The solid lines are the sensitivity derived in bins that contain enough background events (>10) to make a reliable sensitivity calculation. The dashed lines are obtained with a extrapolated background spectrum for bins where there are too few background events. The solid black line is the expected CTA southern sensitivity.

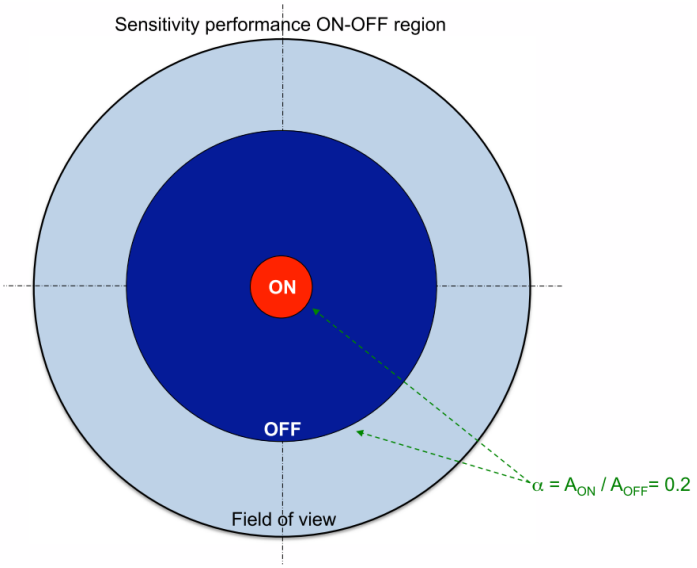


Figure 4.26: Method used in this work of defining the region for background estimation. The ON region from which the signal is determined is based on the angular resolution of the instrument as calculated in 4.8.3. The background level is estimated from an area 5 times larger and surrounding the ON region. Figure from [105].

In both equation 4.8.11 and 4.8.12, E is the energy in TeV [CTA MC Workgroup]. From this the rate of observed events can be calculated as

$$R_{trig} = \frac{N_{trig}}{N_{sim}} \; A \; \Omega \frac{\phi}{\gamma - 1} \; (E_{min}^{1-\gamma} - E_{max}^{1-\gamma}) \tag{4.8.13}$$

where N_{trig} is the number of triggered events passing all cuts, N_{sim} is the number of simulated events, A is the area over which the events where simulated, Ω is the solid angle which defines the area of the ON (gamma) or OFF (proton) regions, ϕ the original spectrum (Equation 4.8.11 and 4.8.12), γ is the spectral index of ϕ and finally E_{min} and E_{max} are the energy bounds of interest. The observed number of events is then simply obtained by multiplying the rate by the observation period and the excess signal events calculated from

$$N_{\text{excess}} = N_{\text{ON}} - \alpha N_{\text{OFF}}, \tag{4.8.14}$$

and then the differential sensitivity is simply in the form

$$\frac{dN}{dE} = \frac{N(E)_{\text{excess}}}{A(E)_{eff} \Delta E \Delta T}, \tag{4.8.15}$$

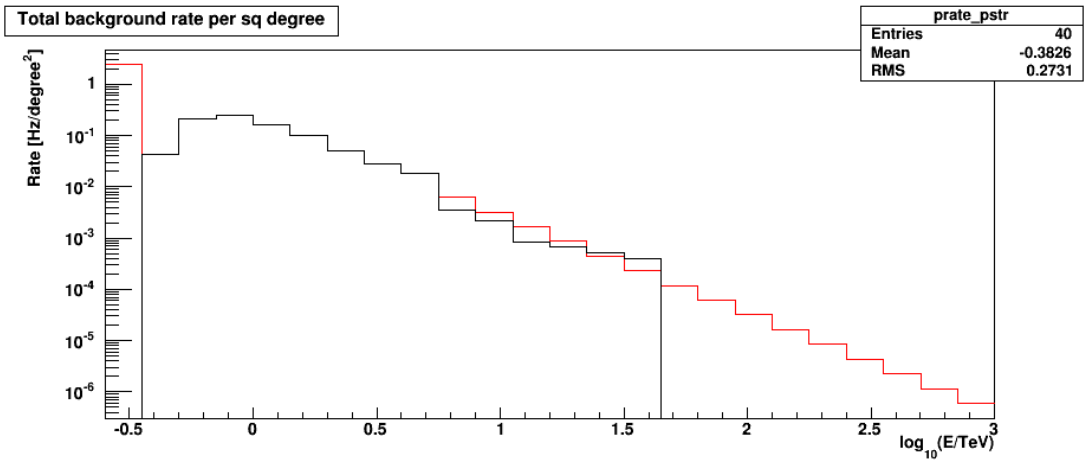


Figure 4.27: Extrapolated proton rate obtained by applying a fit to the lower energy bins and used to estimate the rate at the highest energies. Where the black histogram is the simulated data and the red line is the fit.

where $A(E)_{eff}$ is the signal effective area as seen in Figure 4.23, ΔE is the energy range and ΔT is the observation period.

One problem that often arises in work like this is that the remaining number of protons, especially at high energies, after background rejection and cuts, is too low to reliably calculate the significance. In general, the significance equation from [85] only holds true if the number of events is "not too few", which is taken to be over 10 events. With the mini-array production, it was often the case for higher energy bins that $N < 10$ for protons and therefore in order to have an estimate of performance at high energies, an extrapolation process was adopted. For each set of analyses, a simple power law fit was applied to the proton rate passing all cuts for bins with enough signal events, assigning values from the fit to those bins with less than 10 events. An example can be seen in Figure 4.27). The extrapolated bins are not to be used as reliable estimators of performance and only exist to "guide the eye" until more data can be produced.

The resulting sensitivity calculated for each array spacing can be seen in Figure 4.25, where bins that use the extrapolated proton rate are marked with dashed lines. It can be seen that the general overall trend that has been found in the previous performance estimators unsurprisingly also holds true here, i.e. larger separation

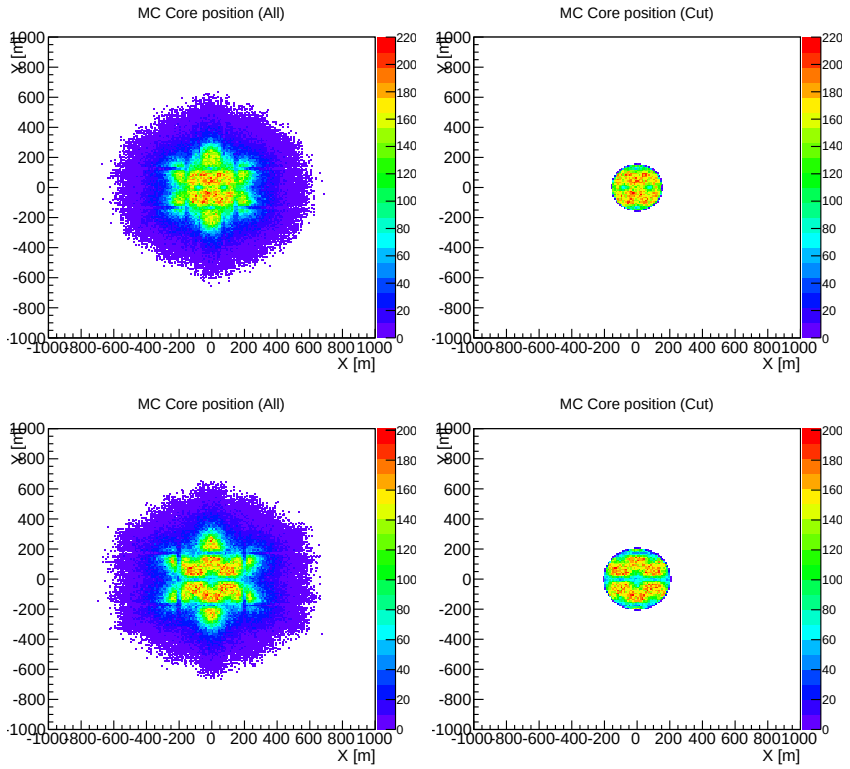


Figure 4.28: True core position for triggered events for 150 m spacing (top) and 200 m spacing (bottom). The left plots shows all triggered events, whereas the right plots show events remaining after a core distance cut equal to the radius of the array has been applied. The colour scale represents the density of events.

improves performance at higher energies.

This gives us an indication of the performance that an initial array of GCT telescopes could achieve in the preproduction phase of CTA. However, the final performance of a larger array will obviously differ due to the increased number of telescopes. Even with a small sample data set it is still possible to obtain an insight into CTA as a whole by considering the fact, that for the full observatory, the events observed are vastly more likely to be contained within the array.

4.9 High Level Evaluation: CTA Sub-Array

In the previous section, the performance of a stand alone mini-array of 7 telescopes was evaluated. Here, an intermediate step towards scaling to the full array is performed. For the southern CTA array, it is envisaged that there will be 70 small

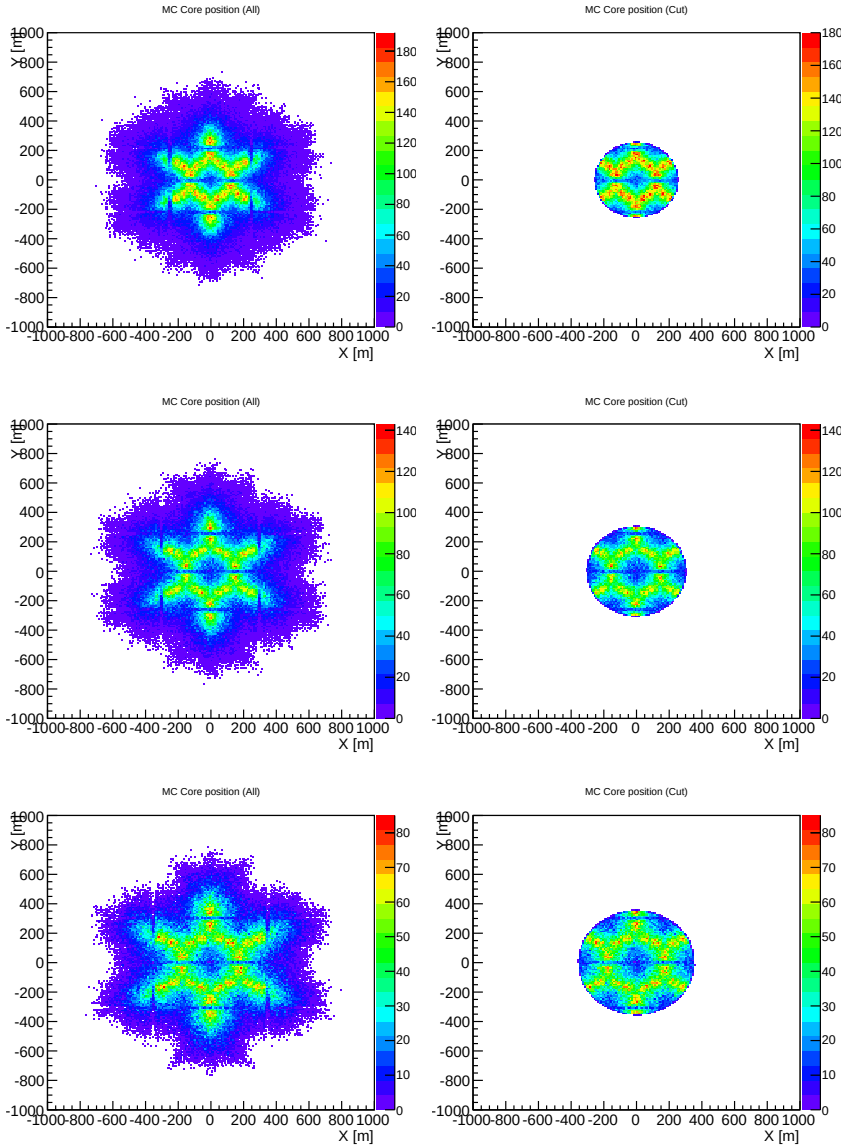


Figure 4.29: Same as Figure 4.28 but for 250 m spacing (top), 300 m spacing (middle) and 350 m spacing (bottom).

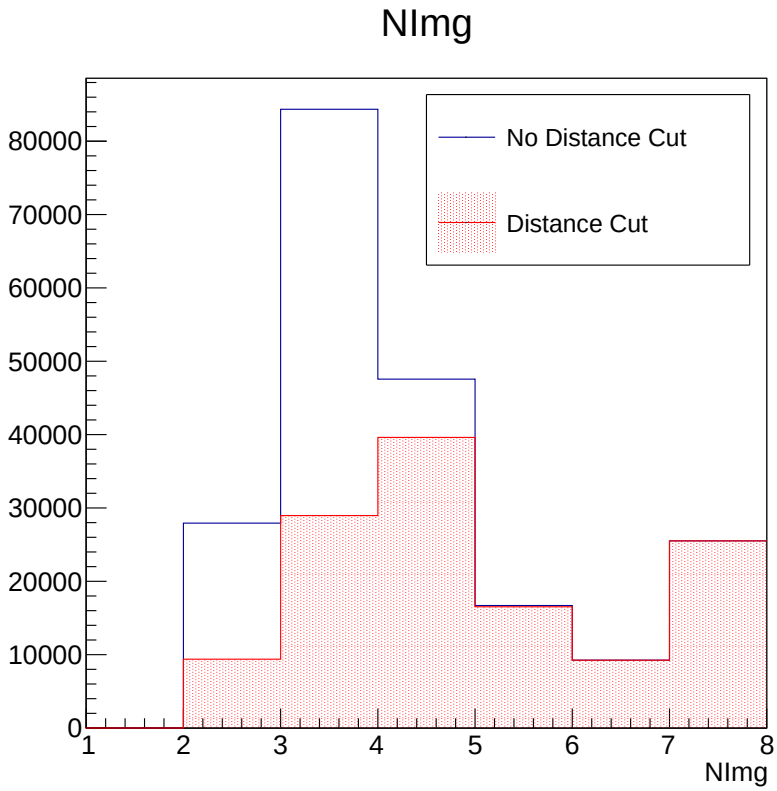


Figure 4.30: Example number of images used in event reconstruction over the whole energy range before (blue) and after the core distance cut (red). The cut effectively removes events with fewer images and therefore events that are more likely to have a worse reconstruction.

size telescopes covering an area of several square kilometres. Therefore the observed events will be detected mainly within the array itself and by performing a cut on impact distance, the 7 telescope mini-array can be evaluated as a sub-array of CTA.

In Figures 4.28 and 4.29 the true core positions of the triggered events for each of 5 arrays with increasing spacing can be seen. In the left column, the full distribution is shown, with an increasing number of events being recorded between the telescope positions (recall the array layouts presented in Figure 4.19). In the right column, a cut on the true core position that removes events which land outside of the array has been implemented. These are for events triggered by GCTM and the horizontal and vertical gaps arise from the gaps within the camera, which reduce the probability of triggering and accurately reconstructing these events. As GCTS has smaller gaps between modules, this effect is less noticeable.

The major effect that a distance cut will have on the analysis concerns the number of images that will be used to reconstruct the shower. In Figure 4.30 the distribution of the number of images included in the stereoscopic trigger before and after the distance cut can be seen. There is a considerable reduction in the number of events imaged by only two or three cameras and therefore a drastic improvement in the angular and energy resolution is expected. In Figures 4.31 and 4.32 it can be seen that the resolutions have indeed been improved overall, with the greatest increase in the angular resolution at around 1 TeV. For the energy resolution, the main improvement is observed at lower energies for the more compact arrays.

In terms of overall sensitivity, most noticeably the more compact arrays see an improved performance below 10 TeV. As this study represents an investigation into a sub-array of the southern CTA observatory, a simple scaling has been attempted in order to represent the larger array. In Figure 4.33, the 7 telescope sub-array has been duplicated 18 times to make up an array of 73 small size telescopes which should result in an increase of the effective area by a factor of 18. Considering Equation 4.8.15 it is evident that the differential sensitivity should scale in a similar fashion. This has been applied and the result can be seen in Figure 4.34 for the effective area and Figure 4.35 where the sensitivity of some of the arrays become comparable to the total CTA southern observatory. It is important to note here

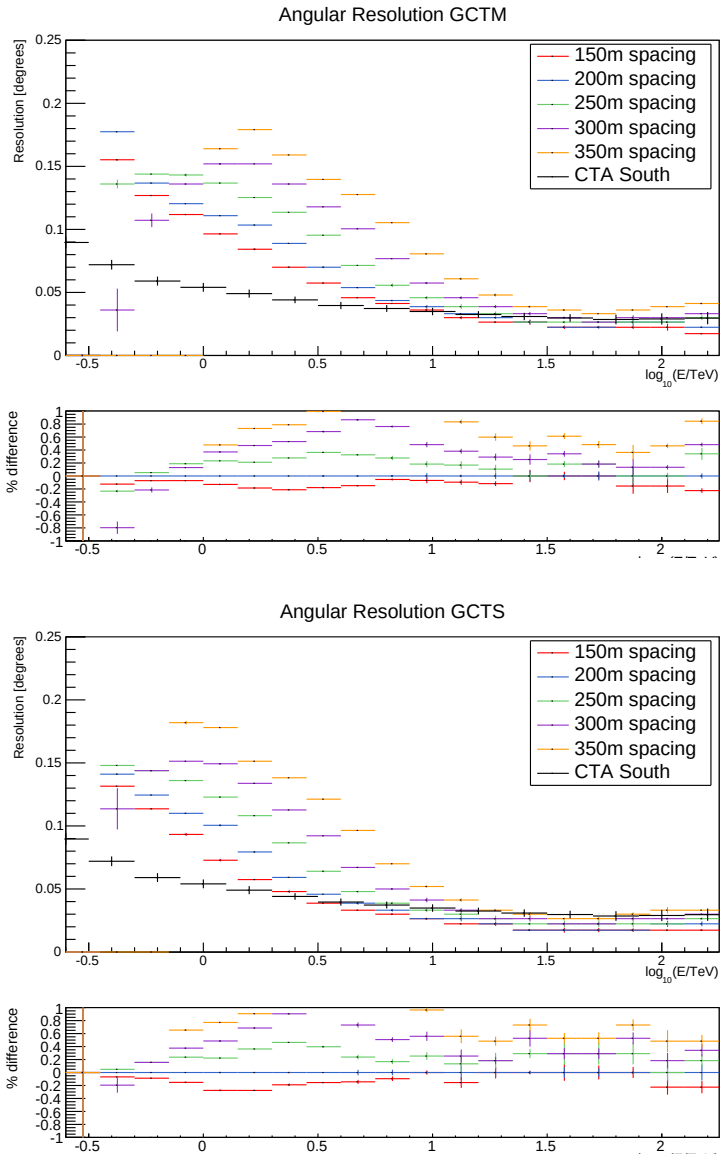


Figure 4.31: Angular resolution obtained for each sub-array for GCTM (top) and GCTS (bottom).

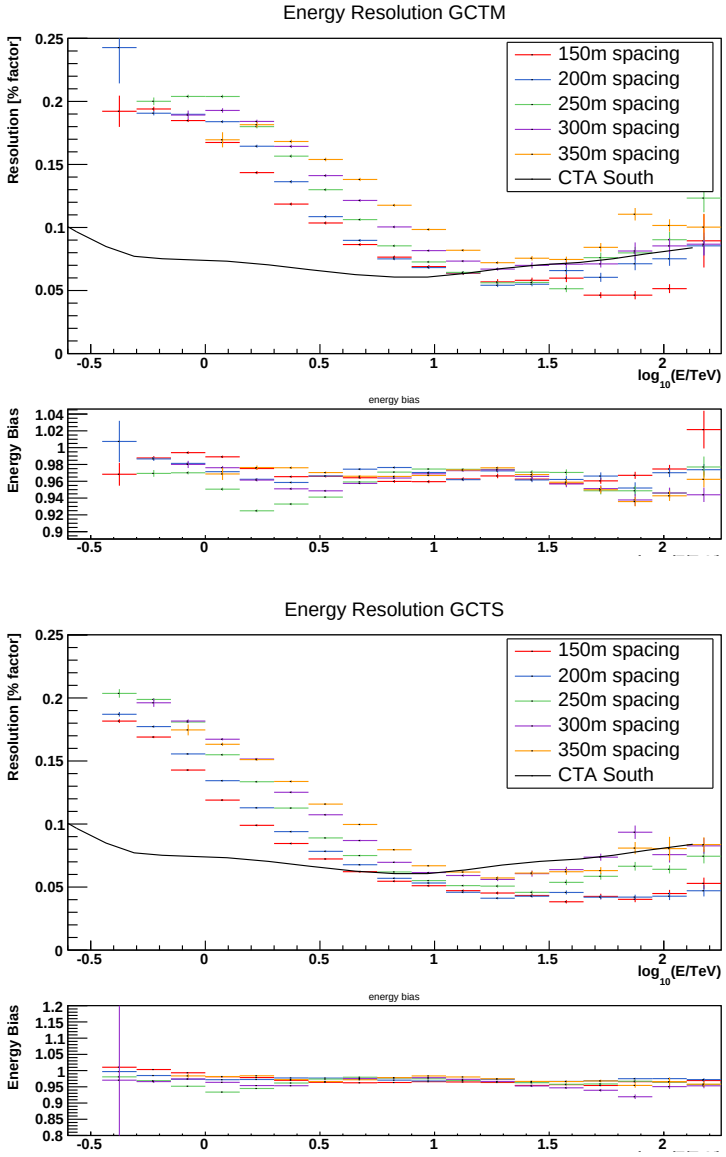


Figure 4.32: Energy resolution and bias obtained for each sub-array for GCTM (top) and GCTS (bottom).

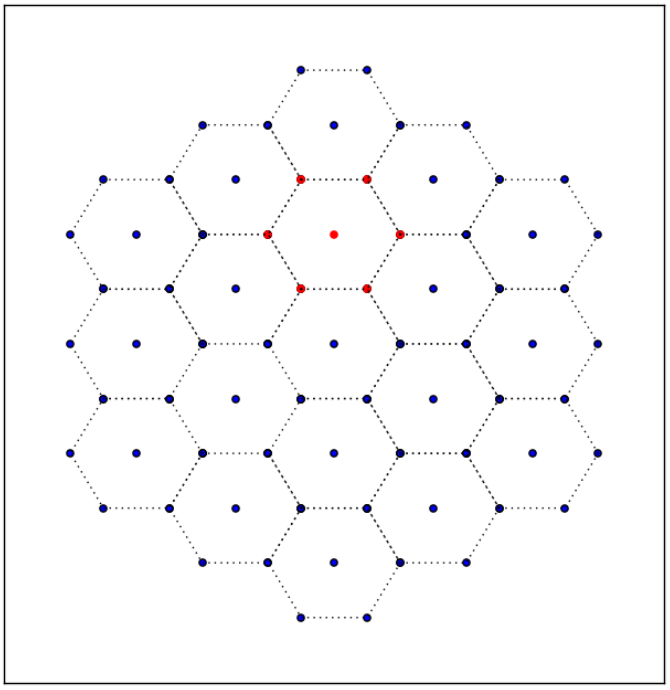


Figure 4.33: Extrapolated array of 18 sub array like sectors consisting of 73 telescopes in total.

that this is performed as a sanity check in order to determine whether the results are within an order of magnitude of the predicted southern array performance.

In order to enable a direct comparison between GCTM and GCTS, a single spacing has been chosen, balancing the improved performance above 10 TeV and the loss below with increasing inter-telescope separation. A spacing of 250 m was chosen and the resulting performance of the two configurations can be seen in Figure 4.36. From this it can be seen that, for the cuts applied, that GCTS out performs GCTM at all energies by around 20% in energy and angular resolution and sensitivity. However, it must be noted that this should not be taken as a definitive answer. In this work both configurations have been analysed in parallel, whereas it may be the case that each would benefit in some way from different optimisations.

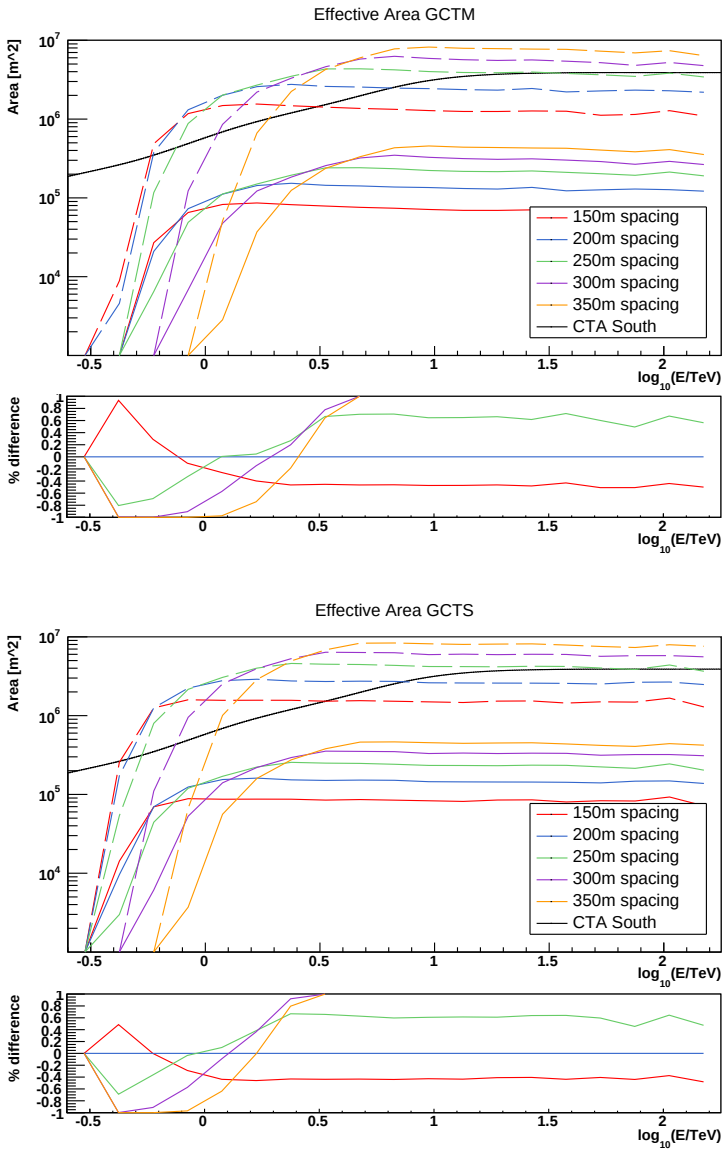


Figure 4.34: Effective area obtained for each sub-array for GCTM (top) and GCTS (bottom).

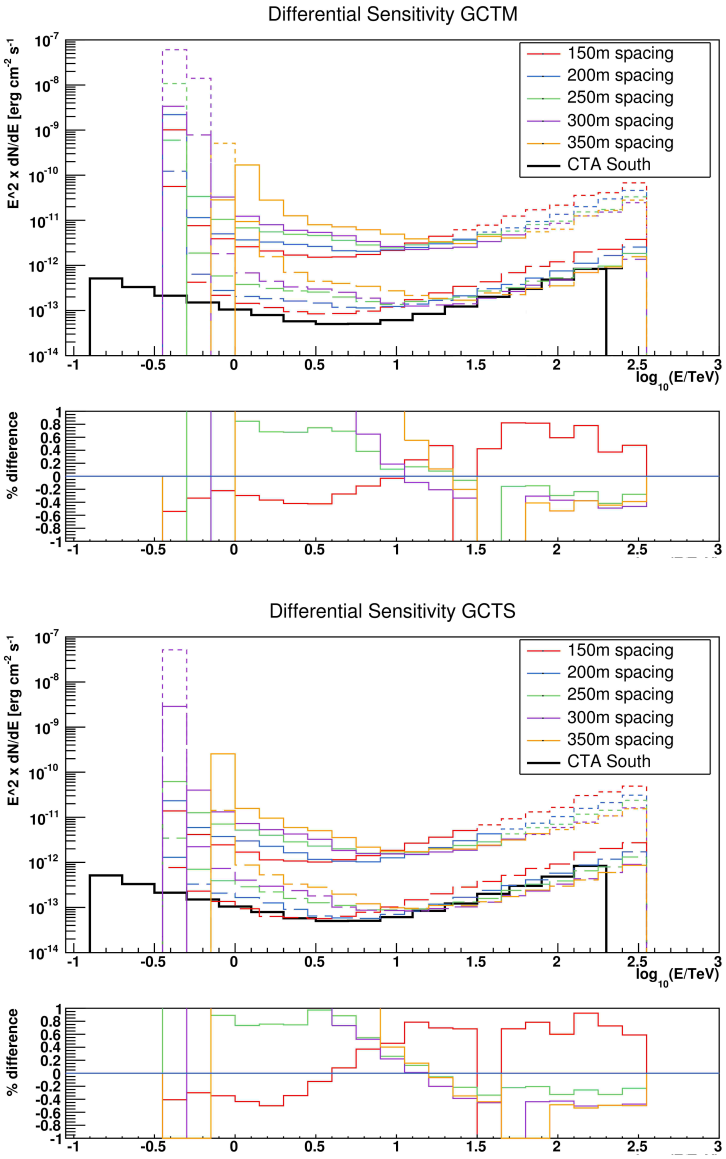


Figure 4.35: Differential sensitivity derived for each array spacing for GCTM (top) and GCTS (bottom). The solid lines are for sub-array sensitivities derived from bins with enough background events while the small dashed lines used an extrapolated background rate. The larger dashed lines are for an extrapolated larger array consisting of 18 sub-arrays, bringing the sensitivity to a level comparable with the CTA south results.

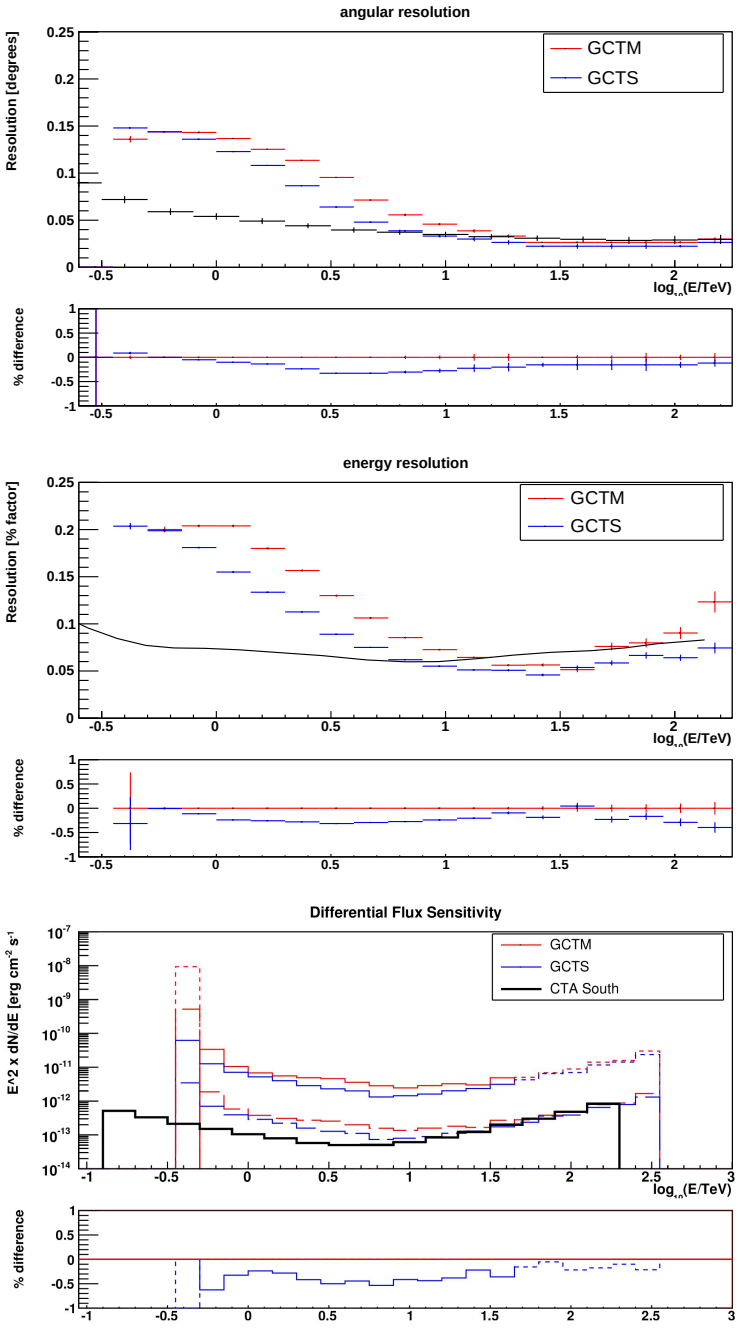


Figure 4.36: Directly comparing the Angular and Energy resolution and the Differential sensitivity for GCTM and GCTS at 250 m separation. For the differential sensitivity an extrapolated 73 telescope array has also been shown (long dashed line).

		GCTM				
Energy	Spacing	A_{eff}	E_{res}	θ_{68}	Diff Sens	Scaled Diff Sens
		m^2	%	deg	erg cm^{-2}s^{-1}	erg cm^{-2}s^{-1}
1 TeV	150 m	82515	**16.74**	**0.096**	**2.59**$\times10^{-12}$	**1.44**$\times10^{-13}$
	200 m	111411	18.39	0.111	3.68$\times10^{-12}$	2.04$\times10^{-13}$
	250 m	**112089**	20.39	0.137	6.80$\times10^{-12}$	3.78$\times10^{-13}$
	300 m	47560	19.28	0.152	12.28$\times10^{-12}$	6.81$\times10^{-13}$
	350 m	2833	16.96	0.164	16.89$\times10^{-12}$	93.84$\times10^{-13}$
	CTA Req		10.0	0.0625		1.88$\times10^{13}$
	CTA Goal		5.5	0.0338		0.86$\times10^{13}$
10 TeV	150 m	71278	6.90	**0.036**	**2.14**$\times10^{-12}$	**1.19**$\times10^{-13}$
	200 m	135340	**6.82**	0.039	2.21$\times10^{-12}$	1.23$\times10^{-13}$
	250 m	222810	7.27	0.046	2.45$\times10^{-12}$	1.36$\times10^{-13}$
	300 m	327373	8.17	0.057	2.62$\times10^{-12}$	1.46$\times10^{-13}$
	350 m	**456030**	9.84	0.081	4.94$\times10^{-12}$	2.74$\times10^{-13}$
	CTA Req		10.0	0.0372		1.51$\times10^{13}$
	CTA Goal		5.0	0.0189		0.83$\times10^{13}$
100 TeV	150 m	71097	**5.15**	**0.022**	21.64$\times10^{-12}$	12.02$\times10^{-13}$
	200 m	127138	7.52	0.027	13.58$\times10^{-12}$	7.54$\times10^{-13}$
	250 m	212818	9.03	0.027	9.56$\times10^{-12}$	5.31$\times10^{-13}$
	300 m	290724	8.54	0.030	8.58$\times10^{-12}$	4.77$\times10^{-13}$
	350 m	**410105**	10.16	0.039	**6.34**$\times10^{-12}$	**3.52**$\times10^{-13}$
	CTA Req		10.0	0.0255		6.26$\times10^{13}$
	CTA Goal		5.0	0.0125		3.98$\times10^{13}$

Table 4.3: A summary of the results for each array spacing using GCTM telescopes including a distance cut, comparing results at 1 TeV, 10 TeV and 100 TeV. Where A_{eff} is the effective area, E_{res} is the energy resolution, θ_{68} is the angular resolution, Diff Sens is the differential sensitivity and the Scaled Diff Sens is the sensitivity for an extrapolated array of 73 telescopes.

GCTS

Energy	Spacing	A_{eff}	E_{res}	θ_{68}	Diff Sens	Scaled Diff Sens
		m^2	%	deg	erg cm^{-2}s^{-1}	erg cm^{-2}s^{-1}
1 TeV	150 m	86907	**11.90**	**0.0728**	**1.69**$\times10^{-12}$	**0.94**$\times10^{-13}$
	200 m	154542	13.44	0.1005	2.99$\times10^{-12}$	1.66$\times10^{-13}$
	250 m	**169827**	15.50	0.1229	5.19$\times10^{-12}$	2.88$\times10^{-13}$
	300 m	140314	16.73	0.1493	7.27$\times10^{-12}$	4.04$\times10^{-13}$
	350 m	55968	16.32	0.1780	15.72$\times10^{-12}$	8.73$\times10^{-13}$
	CTA Req		10.0	0.0625		1.88$\times10^{13}$
	CTA Goal		5.5	0.0338		0.86$\times10^{13}$
10 TeV	150 m	83080	**5.11**	**0.0265**	1.85$\times10^{-12}$	1.03$\times10^{-13}$
	200 m	145106	5.32	**0.0265**	**1.26**$\times10^{-12}$	**0.70**$\times10^{-13}$
	250 m	235248	5.51	0.0332	1.59$\times10^{-12}$	0.89$\times10^{-13}$
	300 m	330486	6.15	0.0412	1.54$\times10^{-12}$	0.86$\times10^{-13}$
	350 m	**453713**	6.69	0.0520	1.77$\times10^{-12}$	0.99$\times10^{-13}$
	CTA Req		10.0	0.0372		1.51$\times10^{13}$
	CTA Goal		5.0	0.0189		0.83$\times10^{13}$
100 TeV	150 m	92747	**4.49**	**0.0173**	16.59$\times10^{-12}$	9.21$\times10^{-13}$
	200 m	148407	4.27	0.0224	10.37$\times10^{-12}$	5.76$\times10^{-13}$
	250 m	246914	6.41	0.0224	6.99$\times10^{-12}$	3.88$\times10^{-13}$
	300 m	320910	7.57	0.0265	5.93$\times10^{-12}$	3.29$\times10^{-13}$
	350 m	**441304**	8.05	0.0332	**5.34**$\times10^{-12}$	**2.97**$\times10^{-13}$
	CTA Req		10.0	0.0255		6.26$\times10^{13}$
	CTA Goal		5.0	0.0125		3.98$\times10^{13}$

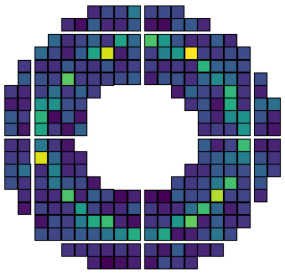
Table 4.4: Same as Table 4.3 but for GCTS.

4.10 Conclusion

In this Chapter an overview of GCT as represented within a simulation framework has been given, followed by a detailed evaluation of its performance. It is apparent that both the MaPM camera based telescope, GCTM, and the SiPM camera based GCTS, will meet the required performance as defined by the CTA consortium, at least in terms of trigger efficiency and charge resolution. In an investigation into inter-telescope spacing, ranging from 150 m to 350 m, results were presented for both GCTM and GCTS as part of an initial stand-alone array and as part of a larger array of telescopes, representing the full SST component of CTA. In all of these a clear trend is apparent for the sub-arrays in which a larger spacing results in an improved sensitivity at higher energies, but there is an improved performance in terms of angular and energy resolutions for smaller spacing. As the choice of telescope spacing will be based on an optimisation of performance for the full array, it is the sub-array performance which is of interest here, with the mini-array results providing an insight into the capability of possible early observations during construction. Therefore, a summary of results for GCTM and GCTS at energies of 1 TeV, 10 TeV and 100 TeV for each spacing for the sub array is presented in Tables 4.3 and 4.4. Although it would be desirable to conclude an optimal spacing from these results, it is clear a choice must be made including the performance of the MST and LST component of the array while also taking into consideration the desired science outcomes of CTA. In addition, the results presented here have not been fully optimised and in order to fully understand the performance at the highest energies, as is seen by the required extrapolation method, more background data need to be produced. The optimisation and additional data production will represent future work within the GCT Monte Carlo group.

Chapter 5

Muon Calibration for GCT



The following chapter discusses the use of unaccompanied muon showers incident along the mirror axis as a means of performing absolute calibration of the telescope efficiency the GCT. Due to the nature of locally-produced muon showers, the image recorded on the camera focal plane is in the form of a ring. By reconstructing the momentum, direction and impact point the amount of expected light can be estimated and compared to the observed value, arriving at an idea of the overall efficiency of the optical throughput of the system. Within CTA, nestled under the working group focused on common camera calibration facilities, there is a group representing each telescope type that has been working on muon calibration. Discussion within that group has been a main source of direction for this work and a lot of the theory can be found in [69] which in turn is based on [36] (in German).

5.1 Introduction

One of the first measurements that alluded to the presence of a particle that was electromagnetically similar to the electron, but of greater mass, was in 1937 by Neddermeyer and Anderson in their paper *Note on the Nature of Cosmic Ray Particles*. In this, they considered the energy loss of particles occurring in cosmic ray showers by way of a plate of platinum placed across a cloud chamber. They concluded that the most likely explanation for the observation was

> "That there exist particles of unit charge, but with a mass (which may not have a unique value) larger than that of a normal free electron and much smaller than that of a proton. This assumption would also account for the absence of numerous large radiative losses, as well as for the observed ionization." [95]

In the period after this the muon (μ) became the accepted particle responsible and settled into theory. However, it was only in 1991 that the first muon ring was observed in an IACT with the Whipple telescope[1]. In a paper by Fleury and the Whipple collaboration, they reported the "frequent occurrence of ring shaped images" [67]. Later, it became apparent that these muon rings could be a useful source of calibration for Cherenkov telescopes as the light around the ring, along with the radius of the ring, could be directly related to the muon shower and the amount of light that would be expected [118].

In this Chapter the use of these muon rings as a form of calibration for GCT will be investigated. In Section 5.2 some basics about the muons involved will be given followed by a discussion of the motivation for calibration in Section 5.3. In Section 5.4 the spectrum of muons will be described, focusing on the model which is assumed in this work. As mentioned, the Cherenkov light produced by muons results in a ring image in the camera, and the geometry behind this, which will lead

[1] As Whipple was a single dish system, triggers from muon events where a much larger problem. As will be seen later, for telescopes operating in stereoscopic trigger mode these events are not recorded.

to a better understanding of the involved parameters, will be described in Section 5.5. In Section 5.6 the theory that allows for the determination of the efficiency of the system will be given followed by a discussion of the systematics which hinder this process in Section 5.7. The evaluation of this method through MC simulations for GCT will be presented in Section 5.8 with a summary of the results and the determination of the expected rate given in Section 5.9. Finally, as this method needs to work as the optical system degrades, an investigation into this will be presented in Section 5.10.

5.2 Muons

In Section 3.3.2 the secondary particles that are produced in cosmic ray air showers were discussed. The immediate particle of interest here is the lightest known meson, the pion ($\pi^{\pm}$, π^0), and to a smaller extent the kaon (K^+), for it is these that are the main producers of the muon component in an air shower. Recall the proton-proton collision in which each type of pion can be produced:

$$\begin{aligned} p + p &\rightarrow p + n + \pi^+ \\ &\rightarrow p + p + \pi^0 \\ &\rightarrow p + p + \pi^+ + \pi^- \end{aligned} \tag{5.2.1}$$

As shown in equation 2.2.3, the neutral pion π^0 decays by the electromagnetic interaction to produce two photons, which can go on to produce a further electromagnetic cascade. The charged pions ($\pi^{\pm}$) however, decay almost exclusively into a negative muon[2] and an anti-neutrino or a positive(anti) muon and a neutrino,

$$\pi^+ \rightarrow \mu^+ + \nu_\mu, \qquad \pi^- \rightarrow \mu^- + \bar{\nu}_\mu \tag{5.2.2}$$

This process happens early on in the shower development and continues until the the energy of the shower components drop below a certain energy threshold. The

[2]The decay into both electrons and muons undergo helicity suppression, however the suppression factor is much smaller for muons due to their large mass [53]

muons produced have a mean lifetime of 2.2μs, but due to their weakly interacting properties along with relativistic time dilation, they survive down to sea level. It is their larger mass, about 207 times that of the electron, that reduces the effect of electromagnetic interactions such as bremsstrahlung and moves the threshold to higher energies. This leaves ionisation as the primary energy loss mechanism in the atmosphere.

The relativistic muon loses energy via Cherenkov radiation which is emitted at an angle θ_c given by the velocity of the muon $\beta = v/c$ and the refractive index n of the atmosphere,

$$cos\theta_c(x,\lambda) = \frac{1}{\beta(x) \cdot n(\lambda, x)}. \tag{5.2.3}$$

From this the minimum energy needed for a muon to emit Cherenkov radiation can be found by realising that $\beta = v/c = \sqrt{1-(E_o/E)^2}$ and that the phase velocity is required to be greater than the local speed of light ($v > c/n$):

$$E(x,\lambda) > \frac{m_\mu \cdot c^2}{\sqrt{1-1/n(x,\lambda)^2}}. \tag{5.2.4}$$

The lower bound is defined as the threshold energy E_t. Considering only a population of local muons, it can be assumed that the refractive index does not change and therefore the energy threshold for a certain site can be calculated, such as Chile where n=1.00021986 and therefore $E_t \approx 5.04$ GeV.

5.3 Motivation for Calibration

Although muons can be seen as a source of background, the fact that they produce a distinctive ring and contain a known distribution of light makes them a valuable method of calibration that comes at no loss in observation time. In very general terms, muons allow us to understand losses of Cherenkov light in the telescope. The amount of light that makes it through the the system is known as the optical efficiency and it is this that can be monitored in order to improve the understanding of the performance of the telescope as a whole. Additionally, the light that is

observed from muons can be used to further study other aspects of the telescope, as is discussed in the following.

5.3.1 Monitoring Optical Efficiency

The optical efficiency is made up of several contributions:

- **Shadowing:** There will be a constant amount of light that is blocked by the secondary mirror, masts and camera. See Section 4.6.1
- **Mirror Reflectivity:** Depending on the mirror quality, coating and condition the amount of light that is reflected onto the camera may change with time. See Section 4.6.1
- **Acceptance of Detector/Window:** Light may be absorbed by the detector itself, for example by the glass on the front of a MaPM in GCTM, or in the window that protects the SiPMs in GCTS. This will also be affected by the presence of dirt or abrasions. See Section 4.6.2
- **Photon Detection Efficiency:** The detector modules each have a certain efficiency with which they convert Cherenkov photons into p.e. within the camera. See Section 4.6.2

These individual components are combined into a single efficiency,

$$\Psi(\lambda) = \psi_{shadow}(\lambda) \cdot \psi_{mirr}(\lambda) \cdot \psi_{acceptance}(\lambda) \cdot \psi_{PDE}(\lambda) \tag{5.3.5}$$

where $\Psi(\lambda)$ is the total efficiency that is associated with the telescope and determines the amount of light that will be observed from each muon. Periodic measurements of the number of photons observed from muons compared to that which is expected reveals any degradation that might have occurred within the optical system. Such a deterioration is expected to occur over the lifetime of CTA due to ageing/weathering and will be considered further in Section 5.10. Additionally, by considering the efficiency as a function of reconstructed angle, any differences in reflectivity across the mirror can be identified.

5.3.2 Monitoring the Point Spread Function

In a perfect system, the muon would be focused into an infinitely slim ring or arc. However, there is a multitude of effects that act to scatter the light from the muon shower, broadening the resulting image. These can be divided into atmospheric and instrumental effects (see Section 5.7). By fitting the spread of the light around the muon ring, as will be described in Section **??**, the PSF of the optical system can be periodically monitored[3].

5.3.3 Monitoring Flat Fielding

Together with the total light received within the muon ring, the muon light intensity can be calculated on a pixel by pixel basis. The pixel efficiency derived from this should be comparable to the flat fielding obtained from instrumental methods such as the LED flashers [42], part of the calibration system for the GCT camera, a comparison that was also made with the H.E.S.S. telescopes [84].

5.4 Muon Spectrum

The vertical cosmic ray muon spectrum has been measured by a multitude of experiments and is well understood. The most comprehensive parametrisation of the muon spectrum comes from [73], where 25 different data sets of measured atmospheric muon flux above 10 GeV are taken[4] and a theoretical modelling from [40] is applied. The most basic level describes the shape of the spectrum as a function of momentum,

[3] Assuming that the atmospheric effect will be constant over a long enough baseline or is known through measurements

[4] See appendix of [73] for a full description of these data sets

$$\begin{aligned} F(y) =& 10^{H(y)}\mathrm{m}^{-2}\mathrm{sr}^{-1}\mathrm{s}^{-1}\mathrm{GeV}^{-1} \quad \text{with} \\ y =& \log_{10}(p/\mathrm{GeV}) \\ H(y) =& 0.133 \cdot (y^3/2 - 5y^2/2 + 3y) \\ & - 2.521 \cdot (-2y^3/3 + 3y^2 - 10y/3 + 1) \\ & - 5.78 \cdot (y^3/6 - y^2/2 + y/3) \\ & - 2.11 \cdot (y^3/3 - 2y^2 + 11y/3 - 2), \end{aligned} \tag{5.4.6}$$

where $F(y)$ is the flux, $H(y)$ is the spectral index and y is the log of the muon momentum p. There are however several factors that can affect the muon spectrum which will be discussed.

Geomagnetic Effect

The muon has the same electromagnetic properties as the electron, therefore it is clear that there could be some deflection effect from the Earth's magnetic field. For the simulated site in Chile the magnetic field is defined with the corresponding components:

$$B_x = 21.325\ \mu T \qquad B_z = -8.926\ \mu T, \tag{5.4.7}$$

where B_x is the magnetic field strength in the direction of magnetic north and B_z is the vertical component. In CORSIKA, the angle of deflection, defined as α, of a particle with momentum p in the magnetic field B is given by:

$$\alpha \approx lZ\frac{\overrightarrow{p} \times \overrightarrow{B}}{p^2}, \tag{5.4.8}$$

where l is the path length and Z is the charge of the particle. As muons are about 200 times more massive than electrons, the deflection effect is largely reduced. This is confirmed by simulating a set of on-axis muon events and measuring any deviations in the camera image. For the magnetic field strength defined here, the effect seen is negligible, consisting of a maximum deviation of around 0.15° (roughly a pixel width) at the lowest energies (5-10 GeV).

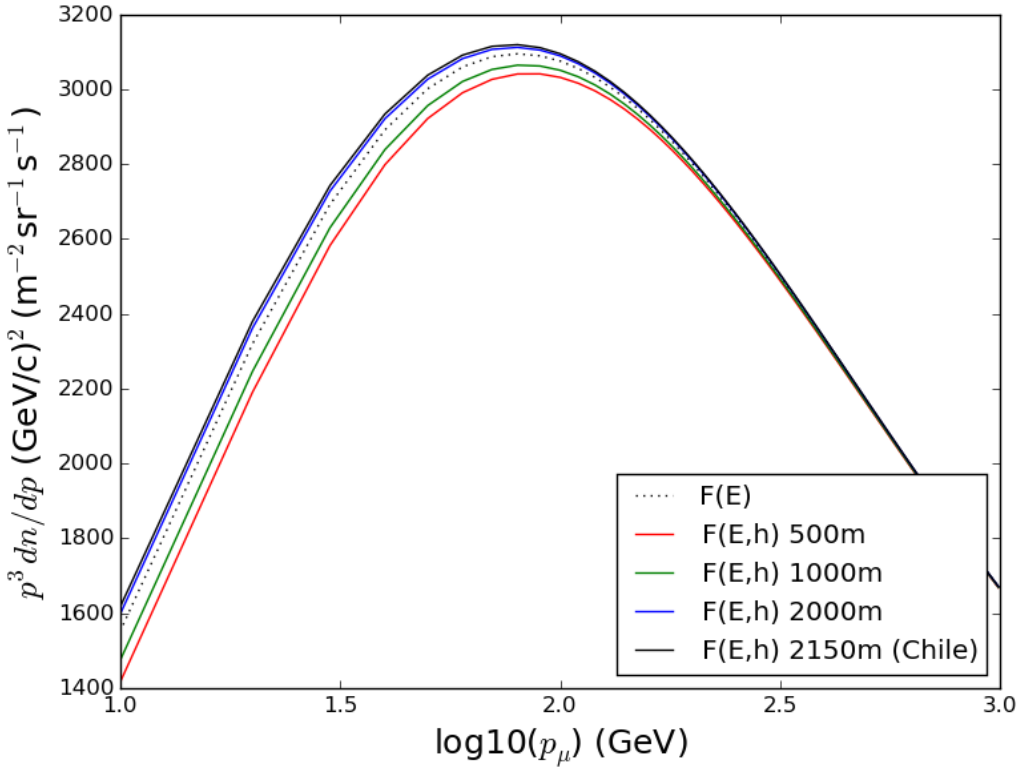


Figure 5.1: Theoretical Flux for a selection of example altitudes where the shape is based on equation 5.4.6 and scaled with height according to equation 5.4.9. F(E) represents the unscaled flux.

Altitude

The altitude at which the muons are observed will have an effect on the measured flux. In this work, the correction factor which is applied to Equation 5.4.6 is taken from [73] and can be expressed as

$$\begin{aligned} F(h) &= F(0) \cdot \exp(h/L) \quad \text{with} \\ L &= 4900\text{m} + 750\text{m} \cdot p/\text{GeV} \end{aligned} \tag{5.4.9}$$

where h is the altitude of the observatory. Several examples can be seen in Figure 5.1. In addition to the selection of the observation altitude, as the muons are assumed to have originated within a hadronic air shower, the maximum starting height must be defined. In order to simulate events that will be visible to the telescope, the maximum height is set in such a way that a muon impacting on the edge of the mirror of radius R will be contained within an image,

$$h_{\text{max}} = 2R \cdot \cot(\theta_c) \ . \tag{5.4.10}$$

therefore resulting in an injection height of $h_{\text{inj}} = h_{\text{obs}} + h_{\text{max}}$. In CORSIKA the maximum starting height is set by the parameter FIXCHI which is defined in g/cm^2 mass overburden. This is calculated using the code available with the CORSIKA_SIMTELARRAY software package EVAL_ATMO which takes the atmospheric profile and maximum radius on the ground, in this case the diameter of the primary mirror, and returns the FIXCHI parameter. For a GCT in Chile this corresponds to 758.782 g/cm^2, or a maximum height of 2.579 km a.s.l (429 m above the telescope).

Zenith Angle

It is expected that there will be a difference in the muon spectrum depending on the zenith angle at which the measurements are being taken. In most published muon spectra data sets, only vertical muons are considered [73]. It is foreseen that the observing strategy for CTA will require that muon calibration data be taken during normal operation and therefore at a range of zenith angles. This effect will be considered in future work once it is shown that in the simplest on axis mode, the muon calibration method is feasible for all telescope types.

5.5 Muon Air Shower Geometry

In an electromagnetic air shower, the resulting Cherenkov radiation is a combination of many emitting particle tracks that roughly follow the direction of the parent particle. In the case of hadronic showers, due to the inelastic scattering involved in the collision of a proton with an atmospheric nucleus, the resulting particles have a large enough transverse momentum to escape the main component of the shower. Therefore it is possible to observe unaccompanied muons and their resulting cone of Cherenkov radiation. To better understand the muon calibration method, it is important to consider the geometry of the muon and telescope, and how these result in a ring in the focal plane.

For simplicity, let us consider an on-axis muon observed by a single mirror telescope. In Section 5.1 it was shown that, if the muon is above a given energy threshold, it will emit a parallel beam of Cherenkov light at an opening angle θ_C. Assuming the track is short, the energy loss due to ionisation and change in the refractive index are negligible and therefore θ_C can be treated as constant. In the case where the true direction of the muon impacts the telescope primary mirror, as in Figure 5.2 a), the light beams emitted at an angle θ_C along the track of the muon which are incident on the mirror will be focused onto the camera focal plane. The light along each radial path from the centre of the mirror to the edge will be focused onto a single point, resulting in a ring of constant magnitude about the centre of the camera.

If a muon impacts the mirror with an impact distance ρ, which is less than the mirror radius, but is travelling parallel to the axis of the telescope, such as in Figure 5.2 b), then the amount of light along each radial path of the mirror is no longer constant. Instead it will vary as a function of ϕ, where the maximum intensity is defined to occur at ϕ_o. This modulation will be discussed further in Section 5.8.6. It should now be easy to see, that in the case where the true direction of the muon does not impact the mirror, as in Figure 5.2 c), the resulting image will be an incomplete ring, consisting of an arc that represents the radial paths that do intersect with the mirror.

As a last consideration, as it is not expected that only on-axis muons will be

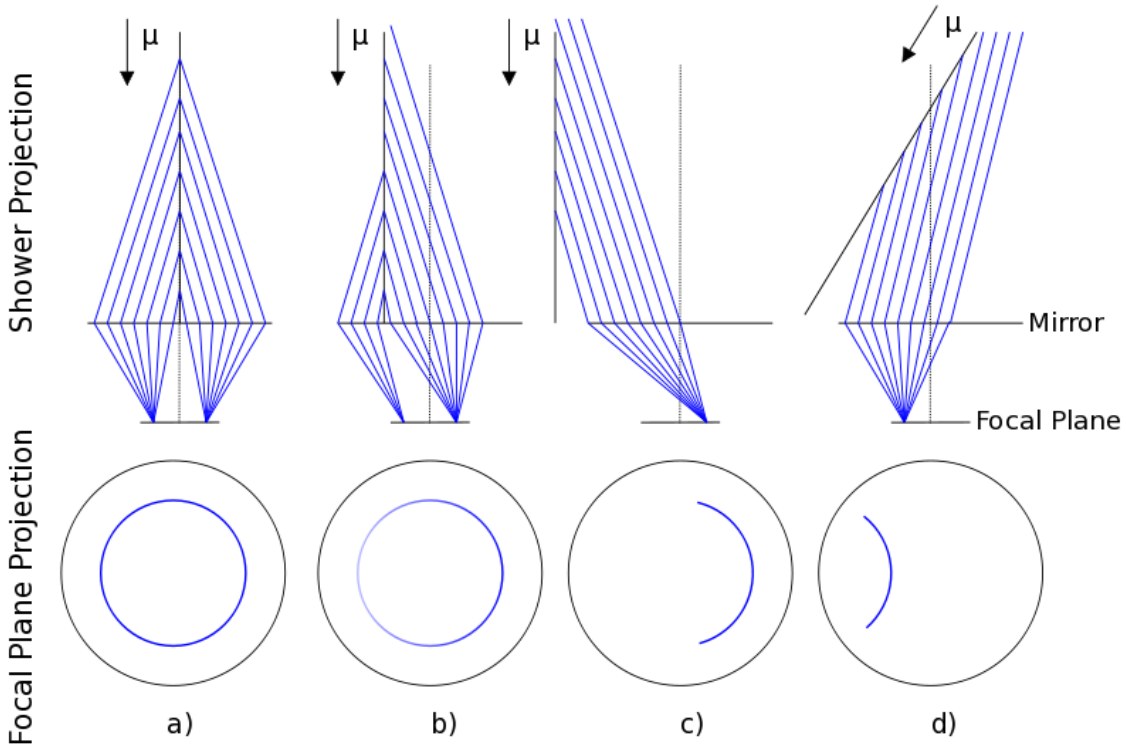


Figure 5.2: Figure showing the muon shower geometry and how the imaged Cherenkov light results in a ring in the focal plane in the simplified case of a single mirror telescope. a) For an on axis muon, the light is radially focused onto a point on the focal plane. Taking into account the rotation about the shower axis this results in an even ring around the centre of the focal plane. b) A muon impacting off axis but within the mirror. In this case some radial paths from the shower core will cover larger sections of the mirror, resulting in a brighter section of the ring in the focal plane. c) Muon landing outside of the mirror; same as b) above but now some radial paths from the muon axis do not intersect with the mirror. d) Muon axis is offset from the telescope axis, resulting in a shift of the ring in the focal plane. Image adapted from [49].

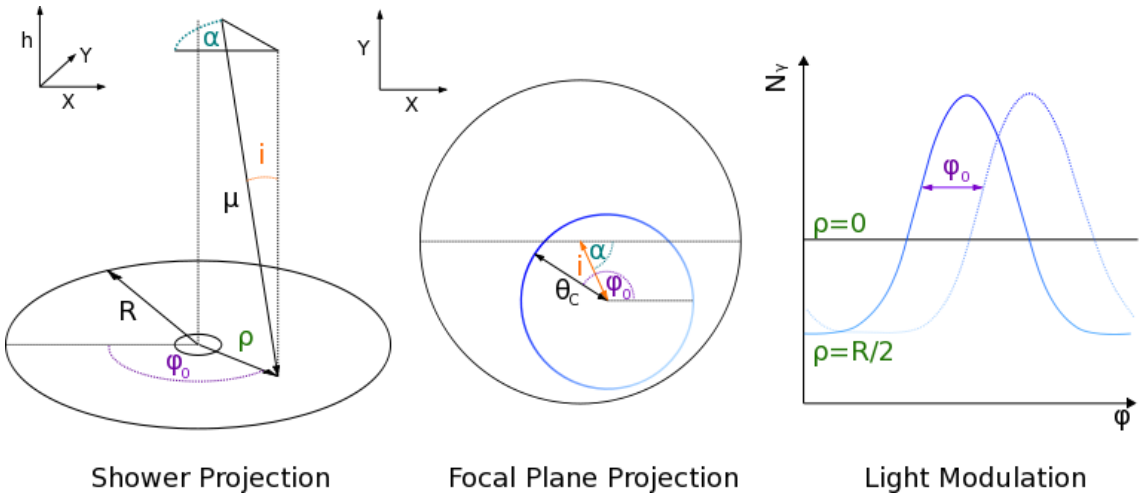


Figure 5.3: Figure adapted from [36] introducing the parameters required for muon analysis. Left: Incident muon geometry where i is the inclination angle from vertical of the muon track and α indicates the direction. R is the radius of the primary reflector, ρ is the impact distance of the muon relative to the centre of the primary and ϕ_0 is the angular offset of the maximum intensity of the ring from the horizontal plane of the camera. Middle: Resulting projection on the camera focal plane, where i, α and ϕ_0 can be obtained from the ring and θ_C is the Cherenkov angle. Right: modulation of the light around the ring from which the shift ϕ_0 in the maximum intensity can be seen and the shape is used to determine the impact distance ρ (See Section 5.8.6).

observed, the effect of an angular offset of the muon's true direction must be understood. In Figure 5.2 d), a muon track intersects with the mirror but at an incident angle of i. In this case the light along each radial path on the mirror is still focused onto a point, but this point will be offset. The total offset of the ring in the camera is given by the angle of incidence of the relativistic muon. A better understanding of the parameters introduced here can be obtained from Figure 5.3, which is adapted from [36].

5.6 Determining Muon Efficiency

Muon calibration relies on the idea that the expected amount of Cherenkov light produced by an unaccompanied relativistic muon can be calculated. In the previous section, the geometry of the muon shower and why this results in a ring was described. In this section, it is now the light within that ring that will be considered, which can be calculated using the Frank-Tamm formula. The number of Cherenkov photons emitted N_{em}, within the wavelength range λ_1 and λ_2, as a function of ϕ, the angle that governs the lateral distribution of the Cherenkov light, and x, the path length of the muon is given by:

$$\frac{d^2N_{em}}{dxd\phi} = \alpha \int_{\lambda_1}^{\lambda_2} (1 - \frac{1}{\beta^2(x)n^2(\lambda, x)}) \frac{1}{\lambda^2} d\lambda \quad \text{photons} \cdot \ \text{m}^{-1} \cdot \text{rad}^{-1}, \tag{5.6.11}$$

where α is the fine structure constant and β is the velocity of the muon divided by speed of light in a vacuum. The equation for the Cherenkov angle has already been presented in Equation 5.2.3; substituting into equation 5.6.11 and assuming the refractive index to be constant[5], the following is obtained:

$$\frac{d^2N_{em}}{dxd\phi} \simeq \alpha \cdot \sin^2\theta_c \cdot I, \tag{5.6.12}$$

where the wavelength dependent component has been abbreviated into a single factor, the assumed Cherenkov spectrum produced from muons:

[5]This is valid if it is assumed that the muon light is produced locally, which is required if we are to observe complete rings.

$$I = \int_{\lambda_1}^{\lambda_2} \frac{1}{\lambda^2} d\lambda. \tag{5.6.13}$$

From here, the total light along the radial length of mirror needs to be calculated, introducing the total path length L along which the muon is visible within the camera. If the muon was incident directly in the centre of the mirror, this would equate to $L = R/\tan\theta_c$. However, if there is any deviation by ρ, the path length will differ from R as a function of the azimuthal angle ϕ. This correction is defined as $D(\rho, \phi - \phi_o)$ where, for a single mirror system this is:

$$D(\rho, \phi - \phi_o) = \begin{cases} 2R\sqrt{1 - (\rho/R)^2 \sin^2(\phi - \phi_o)} & \text{for: } \rho > R \\ R[\sqrt{1 - (\rho/R)^2 \sin^2(\phi - \phi_o)} + \rho/R \cos(\phi - \phi_o)] & \text{for: } \rho \leqslant R \end{cases}$$

The additional corrections needed for a two mirror system will be described in Section 5.8.6. Now, the equation for the number of photons observed can be expressed as,

$$\begin{aligned} \frac{dN_{obs}}{d\phi}(\rho, \phi_o) &\simeq \alpha \cdot \sin^2(\theta_c) \cdot L \cdot I \cdot T \\ &\simeq \frac{\alpha}{2} \cdot \sin(2\theta_c) \cdot D(\rho, \phi - \phi_o) \cdot I \cdot T \end{aligned} \tag{5.6.14}$$

where the important correction factor, T, has also been introduced. This is the transmission of the atmosphere which accounts for the molecular and aerosol extinction of the Cherenkov photons emitted as part of the shower. This has been considered as part of the site studies for CTA and within the muon working group, where a value has been derived for each telescope. In the case of GCT, the weighted transmission is expected to be $T_{total} = 98.9\%$ (See Section 5.7.2).

The final factor that needs to be introduced is the efficiency of the system to observe Cherenkov light from muons,

$$\epsilon_\mu = \int_{\lambda_1}^{\lambda_2} \frac{\Psi(\lambda)}{\lambda^2} \mathrm{d}\lambda \; / \; \int_{\lambda_1}^{\lambda_2} \frac{1}{\lambda^2} \mathrm{d}\lambda \tag{5.6.15}$$

where $\Psi(\lambda)$ is the combined efficiency as expressed in equation 5.3.5. Including this in equation 5.6.14 allows the complete expression for the number of photons observed around the ring to be derived:

$$\frac{dN_{obs}}{d\phi}(\rho,\phi_o) = \frac{\alpha}{2}\cdot\sin(2\theta_c)\cdot D(\rho,\phi-\phi_o)\cdot I\cdot T\cdot\epsilon_\mu. \tag{5.6.16}$$

From this the efficiency of the system for each muon observed can be extracted. It must, however, be made clear that this is the efficiency of the system to detect light from a muon and will not be exactly the same as that for a gamma ray which can be expressed as

$$\epsilon_\gamma = \int_{\lambda_1}^{\lambda_2}\Psi(\lambda)\cdot\frac{T_\gamma(\lambda)}{\lambda^2}\mathrm{d}\lambda \;/\; \int_{\lambda_1}^{\lambda_2}\frac{T_\gamma(\lambda)}{\lambda^2}\mathrm{d}\lambda, \tag{5.6.17}$$

where again $T_\gamma(\lambda)$ is the transmission, which can not be treated as simply as local muons and depends on the initial emission height of the shower particle. The implication of this is that there is a requirement to somehow monitor the conversion factor,

$$C_{\mu\to\gamma} = \frac{\int_{\lambda_1}^{\lambda_2}\Psi(\lambda)\cdot\frac{T_\gamma(\lambda)}{\lambda^2}\mathrm{d}\lambda}{\int_{\lambda_1}^{\lambda_2}\frac{\Psi(\lambda)}{\lambda^2}\mathrm{d}\lambda}\cdot\frac{\int_{\lambda_1}^{\lambda_2}\frac{1}{\lambda^2}\mathrm{d}\lambda}{\int_{\lambda_1}^{\lambda_2}\frac{T_\gamma(\lambda)}{\lambda^2}\mathrm{d}\lambda}, \tag{5.6.18}$$

which would preferably require periodic investigation into the spectral response of the telescope system via the use of, for example, different coloured LEDs in a separate calibration unit. Investigations on this are ongoing and are not considered further in this work. Instead a focus is given to obtaining the muon efficiency for the telescope system.

5.7 Systematic Effects

Having described the theory behind muon calibration, it is now important to revisit the systematic effects that occur and necessitate the need for such a method. In general, these can be divided into two main groups. Those that arise from the instrument itself and those that are a result of the atmosphere, which for Cherenkov telescopes can be thought of as an extension of the telescope.

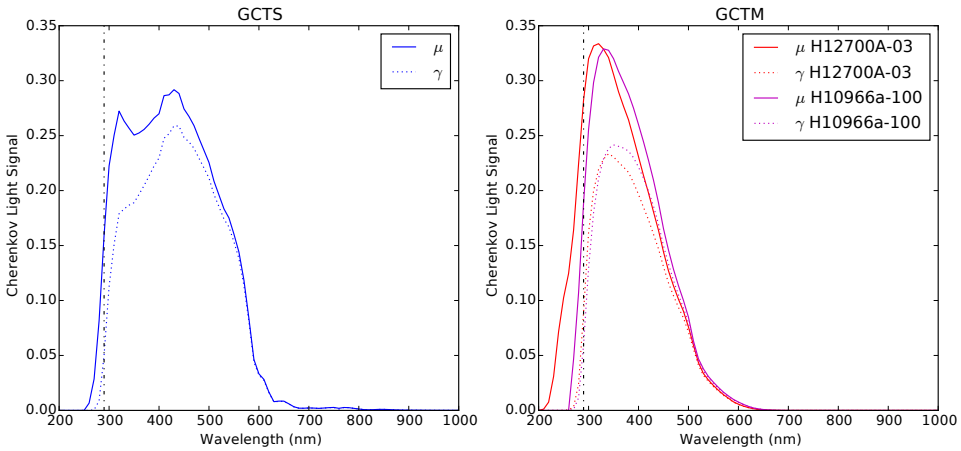


Figure 5.4: The Cherenkov light spectrum from gamma-rays (dashed) and muons (solid) convolved with the optical efficiency of GCTS (left) and GCTM (right). For the gamma-rays, the showers were set to originate from 10 km. In the case of GCTM, two sets of data are shown for two different MaPM devices. The H12700A-03 (current configuration) and H10966A-100 (Similar device but without enhanced UV sensitivity)

.

5.7.1 Instrumentation Effects

Chromatic Degradation

Due to the average emission height and the effect of absorption of Cherenkov radiation from the atmosphere, the spectra from muons and gamma-rays differ, see Figure 5.4. Due to this, if the telescope system undergoes any degradation of efficiency that is wavelength specific, then there would be a possible and potentially unidentifiable change in the muon to gamma efficiency conversion factor $C_{\mu\to\gamma}$. This is has been considered by the CTA muon working group which has derived the requirement that

> "The optical elements of the telescope (mirrors and camera) must be chosen such that the part of the Cherenkov light spectrum from local muons, which stems from wavelengths below 290 nm, must contribute by less than 5% to the observed muon image size, where size is understood as the sum of all photo-electrons contained in the ring image." [B-xST-1500]

The value of 290 nm is chosen as the cut as it is below here that the Cherenkov light from gamma-rays becomes nominal. In Figure 5.4, it can be seen that GCTS

will meet this requirement, as the amount of light below 290 nm makes up only 2.72% of the total. However for GCTM, with the current configuration using the Hamamatsu device H12700A-03 (see Section 4.6.2), there is significant efficiency below 290 nm, contributing 14.66%. This is due to the quoted improved UV sensitivity of that particular device, achieved with an improved glass material. If instead a device with a more traditional window material is considered, such as the H10966A-100 which is also shown in Figure 5.4, the requirement can be met as the total contribution is reduced to 4.5%. It is therefore clear that the choice of detector for GCTM must be considered carefully in order to obtain the optimal results while remaining within requirements.

Primary Shower in image

Muons are produced within cosmic ray showers and therefore the image may contain traces of the parent shower. Additionally, since Cherenkov telescopes nearly always run in stereoscopic mode, in which two or more telescopes are required to trigger, it is likely that in this mode it would be only these contaminated events that are recorded. The muon events that would be required for calibration would not trigger more than one telescope by themselves. To overcome this, some experiments, such as H.E.S.S. periodically perform dedicated muon runs with a mono trigger [49]. However for CTA it is required that each telescope (excluding the LST) must be able to flag potential muons from uncalibrated data with an efficiency of at least 90% above 20 GeV. This would create an effective override to the stereo trigger. Current work by members in other SST telescope groups indicate that this should be achievable. This has yet to be evaluated for GCT but will be investigated in future work.

Pixel Baseline

An accurate estimation of the number of p.e., and therefore the efficiency, from the image requires a good knowledge of the calibrated pixel baseline. This baseline can vary over time and temperature, the latter being even more important for GCTS, as the performance of silicon can be heavily affected by changes in temperature.

Although it is expected that this should be controllable and the pixel baseline will be known to at least a value of 0.2 p.e., it is possible to monitor this baseline with muons. By measuring the off ring pixel intensity, any changes to the baseline should become apparent.

Non-uniformity of Camera and Mirrors

So far only a uniform degradation of the system has been considered, however it is possible that there might arise a case where parts of the camera or mirrors, for example, deteriorate faster than the whole of the system. For example, if one of the six petals in the GCT primarily reflector has a reduced reflectivity compared to the rest, this could affect the reconstruction of the expected number of p.e. observed. However, if the mean efficiency as a function of ϕ_0 is measured, signs of a gradient indicating non-uniformity in the system could be identified.

5.7.2 Atmospheric Effects

Atmospheric Broadening of Ring images

As the muon passes through the atmosphere, it will undergo scattering from molecules and aerosols within the atmosphere. Each interaction causes the muon to lose energy to the atomic potential/charge of the atom responsible for the scattering. Initially it was assumed that a relativistic muon will produce Cherenkov radiation at a constant angle. From equation 5.2.3 it can be seen that if the muon were to lose energy, the Cherenkov angle would decrease. This therefore leads to a blurring of the muon ring as it is eventually seen within the camera. In addition, the muon will lose energy through ionisation which will result in the same effect.

In equation 5.2.3 it can be seen that the Cherenkov angle is also governed by the refractive index. Although the muon showers considered here are produced locally, there will still be some effect of variations of the refractive index which will again add to the broadening of the ring.

These effects, along with the discrete pixel width of the camera and the PSF of the instrument, lead to a final spread of the light along the ring which is defined

as the *Arcwidth*. An attempt to monitor this from data will be discussed in section 5.8.5.

Atmospheric Transmission

A fraction of Cherenkov photons that are produced by the muon will undergo extinction from the atmosphere. The remaining fraction of light is governed by the atmospheric transmission which in turn consists of molecular and aerosol components. This can be very generally expressed as

$$T(r,\lambda) = \exp[-\int_0^r \alpha_{mol}(x,\lambda) + \alpha_{aer}(x,\lambda)dx] \tag{5.7.19}$$

where r is the distance from the telescope, α_{mol} and α_{aer} are the molecular and aerosol extinction factors. This is then integrated over the path length over which the muon emits light in order to find the total transmission.

It has been shown in [69] that the transmission values for GCT at the southern site in Chile, weighted to the Cherenkov spectrum, equate to $T_{mol} = 0.995$ and $T_{aer} = 0.992$ leading to a total atmospheric transmission of 0.987. However, there will be a variation in this factor depending on weather conditions. Therefore, in order to maintain a correct value for the transmission, alternative methods to monitor atmospheric conditions must be employed.

5.8 Evaluation of Method with MC Simulations

In order to study whether GCTM and GCTS will be able to efficiently trigger on "usable" muon rings, such that they can be reconstructed accurately, Monte Carlo simulations of unaccompanied local muon events using CORSIKA have been performed. Using the atmospheric parameters specific for the CTA southern site in Chile located at 2150 m a.s.l., 1×10^6 μ^- between 6 GeV and 300 TeV originating within the camera field of view (half angle of 4.7 degrees) were generated. While the generic equations presented in Section 5.4 for the muon spectrum cover both μ^- and μ^+, it is noted that the differences do not appear within the simulations and therefore it is assumed that the data set can be treated as a sum of μ^- and μ^+.

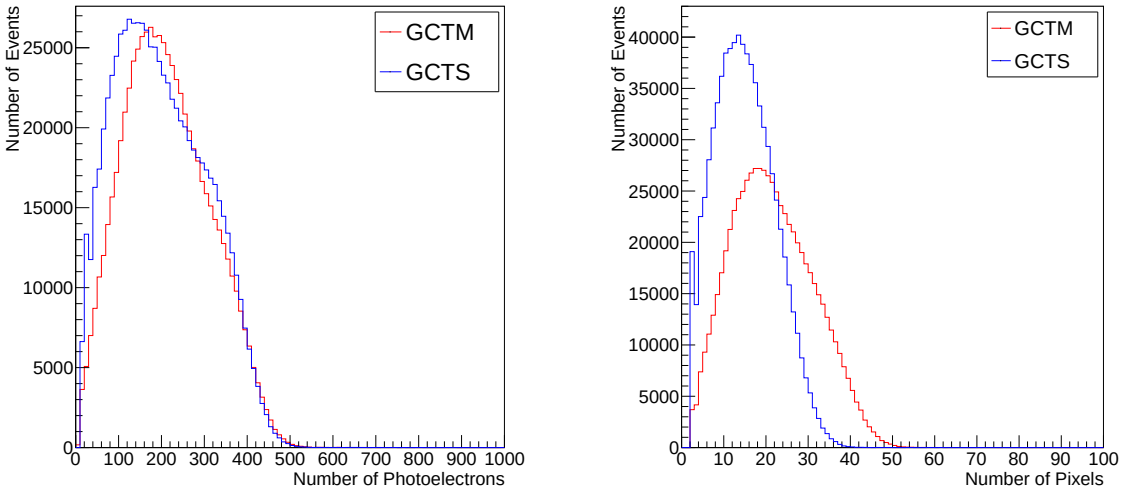


Figure 5.5: Distribution of the number of p.e. and pixels in each image for GCTS. These are what remains after the tail-cuts and are used later in the ring fit.

The generated muon showers were scattered over an area a little over twice the radius of the primary reflector, such that the maximum distance from the telescope would be 4.4 m. The telescope configurations for this chapter are the same as presented in detail in Section 4.6.

5.8.1 Pre-Selection Cuts

In order to reduce the data set and remove events which would end up being unusable, several initial cuts were considered.

Tail Cuts

As a first step, the pixels within the camera that will be part of the muon ring image must be selected. This is done using tail-cuts which add pixels to an image if the pixel has at least n p.e. and has a neighbouring pixel with at least m p.e.. In order to select the most image pixels within the ring, without including any large fluctuations from the NSB, it was found that initial values for GCTM were 3 and 6, while for GCTS which has a larger contribution from NSB, higher cuts of 4 and 8 were required. The values presented here are only best estimates and will need optimising in the future.

Number of photo-electrons and Image Pixels

Previous muon calibration methods have discarded events consisting of only a small arc or with a small radius, as the small number of pixels would result in larger errors on the reconstructed parameters. As an initial choice, a cut of 10 pixels with at least 40 p.e. post tail cuts was chosen. While the p.e. cut may have a minimal effect, the pixel cut removes a large proportion of events that would likely provide poor reconstructions, see Figure 5.5.

5.8.2 Energy Reconstruction, Circle Fitting

Once a reasonable data set has been obtained, the geometric ring that the muon creates within the camera focal plane must be reconstructed. Fortunately the need for fitting circles/arcs has been well studied in many fields and represents one of the most basic tasks in pattern recognition. For the purpose of fitting muon rings, the algebraic algorithm proposed by G. Taubin [114] is used[6].

The basis of the Taubin fit is the minimisation of the equation,

$$\mathcal{F}_T = \frac{\sum_i^n [(x_i - a)^2 + (y_i - b)^2 - R^2]^2}{4n^{-1} \sum_i^n [(x_i - a)^2 + (y_i - b)^2]} \tag{5.8.20}$$

where x_i and y_i are the individual data points, n is the total number of points, and a and b are the x and y coordinates of the centre of the circle and R is its radius.

For the full data set, the fitting algorithm is applied to each image, returning the ring centroid and radius. Here further cuts are applied to obtain better quality rings that are usable for further analysis.

5.8.3 Post Fit Cuts

The cuts based on the geometrical fit to the muon ring have been chosen by the muon working group within CTA. Unlike the previous cuts, these can be seen as telescope independent.

[6]The generic algorithm used can be found at http://uk.mathworks.com/matlabcentral/fileexchange/22678-circle-fit–taubin-method- and was adapted to work within READ_CTA.

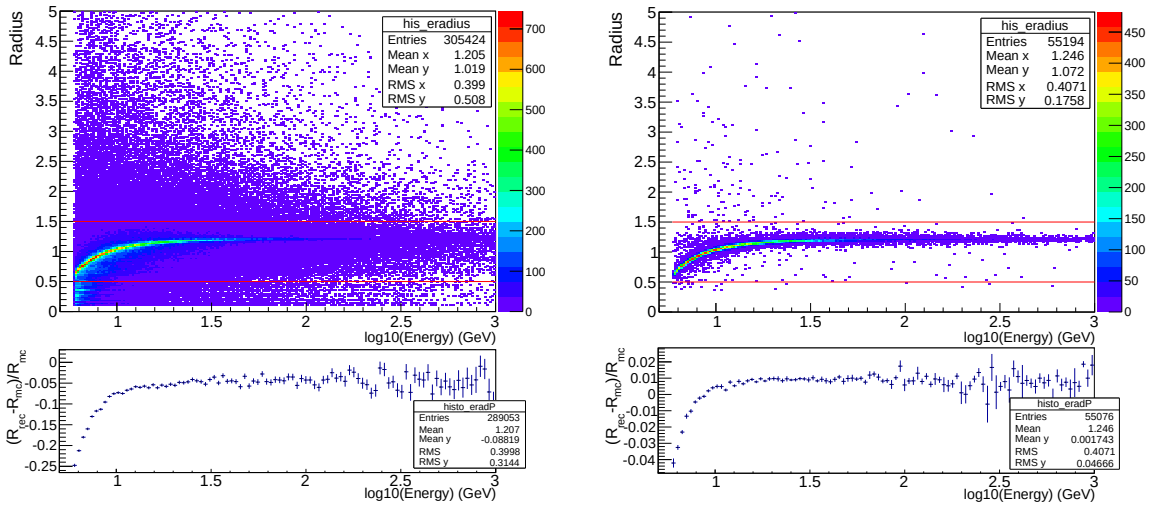


Figure 5.6: Reconstructed radius (top) and resolution (bottom) as a function of energy for GCTS before pixel and p.e. cuts (left) and after cuts (right). On each plot the cuts on the reconstructed radius are also been shown. As can be seen, these are of little significance after the other cuts have been applied.

Reconstructed radius limit

From equation 5.2.3, the expected muon ring radius can be calculated as a function of energy. For the chosen site and the simulated energy range, the minimum Cherenkov angle should be 0.66 degrees at 6 GeV. As the energy the muon increases the Cherenkov angle tends towards a value of 1.201 degrees. Therefore a range of acceptable reconstructed radius values was set to be between 0.5-1.5 degrees. The effect of this can be seen in Figure 5.6, where, without the presence of other cuts, there is a large spread in the reconstructed radius.

Edge cut

To remove undesirable effects arising from artefacts (inefficiencies and aberrations) at the edge of the camera, muons with a reconstructed ring that passes through the outer 0.3° of the camera are rejected[7]. For GCT this corresponds to a condition of,

[7]This is the same for each telescope. The pixel field of view will determine the amount of the camera edge that will be excluded.

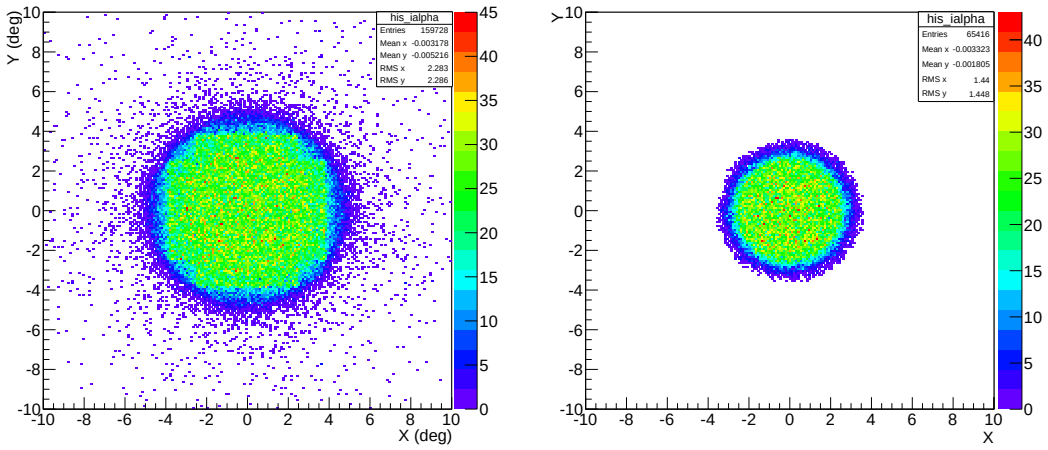


Figure 5.7: Effect of Edge cut: Distribution of events before (left) and after (right) the edge cut.

$$\sqrt{x_c^2 + y_c^2} + R_{reco} < FoV - 0.3^\circ \tag{5.8.21}$$

where (x_c, y_c) are the coordinates of the centre of the ring and R_{reco} is the reconstructed ring radius. This can be seen in Figure 5.7, where the left panel shows the limit of the "usable" camera area and the right panels show the effect of the cut on the distribution of the ring centre coordinates.

5.8.4 Selection Efficiency Post Cuts

In Figure 5.8, the effect of each of the individual cuts that have so far been presented is shown. To better understand the effects of each cut, only events landing within the primary mirror are considered here, as it will be shown in Section 5.8.6 that the events outside will be rejected. It can be seen that the cut on the number of pixels (GCTS) and the edge cut currently remove the largest number of events, with all cuts removing around ∼70% of events.

5.8.5 Ring Broadening, Fitting *Arcwidth*

Once a good fit for the muon ring has been obtained and post fit selections performed, the distribution of light can be investigated. Due to a combination of secondary effects (See Figure 5.11), with the main component coming from mirror

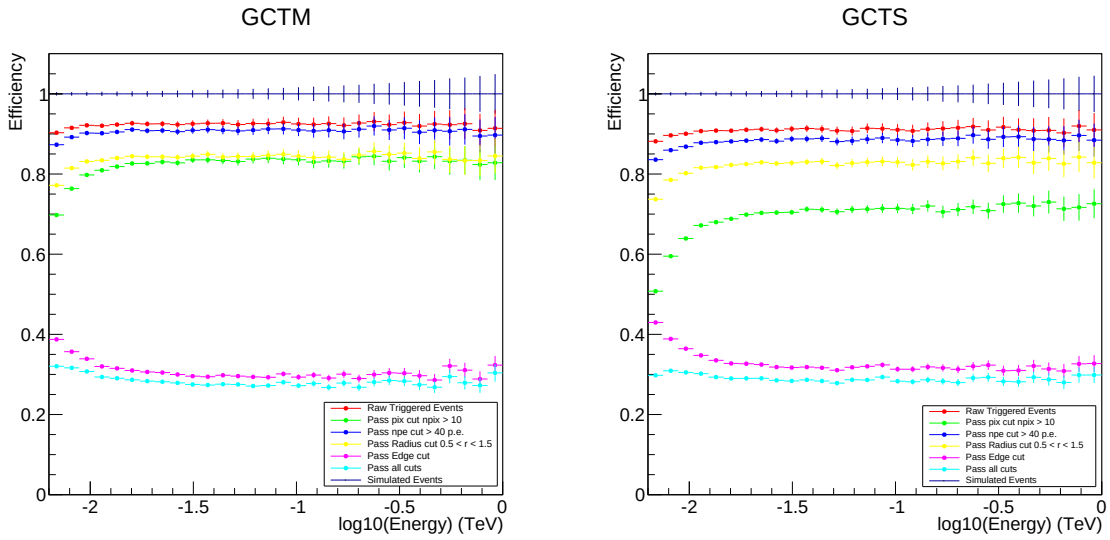


Figure 5.8: Selection Efficiency of muon events for each individual cut presented so far, along with the total selection efficiency.

aberrations, there is a blurring of the muon ring. For GCT, this broadening results in the light being distributed over more than one pixel. In previous experiments such as MAGIC, this broadening effect has been used to monitor the instrument PSF [70]. It is therefore desirable to investigate whether this method would be possible for GCT.

The broadening effect can be seen as a Gaussian smoothing along the radial distribution of the ring, where the parameter *Arcwidth* is now defined as one standard deviation of that Gaussian. In order to obtain the *Arcwidth* for a muon ring, the pixel data within $\pm 0.26°$ of the ring radius is taken. This is then divided into a number of bins depending on the radius of the ring, such that $nbins = 25 \cdot (R_{reco}/1.2)$. The data in each of these bins are then fitted with a simple Gaussian of the form

$$\frac{dN}{dr} = a \cdot \exp^{\frac{-(x-x_0)^2}{2\sigma^2}}, \tag{5.8.22}$$

where a is the amplitude, x_0 is the offset and σ is the *ArcWidth* (see Figure 5.9 for an example). By taking the *Arcwidth* value for each fit within the data set, an indication of the PSF can be derived. An initial view of this can be seen in the top left of Figure 5.10. However, a tighter data selection may produce a more representative indication for the PSF. For example in Figure 5.11 the individual

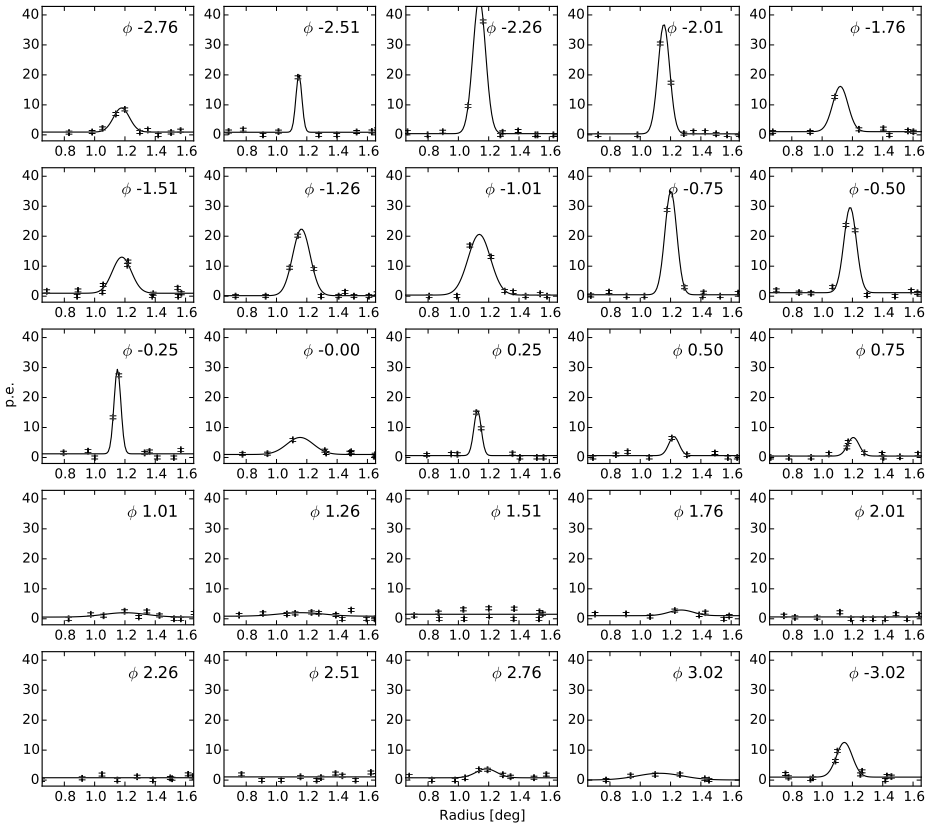


Figure 5.9: Determination of *Arcwidth* from equally spaced ϕ bins around the ring. A Gaussian fit is applied to each bin and the width is recorded.

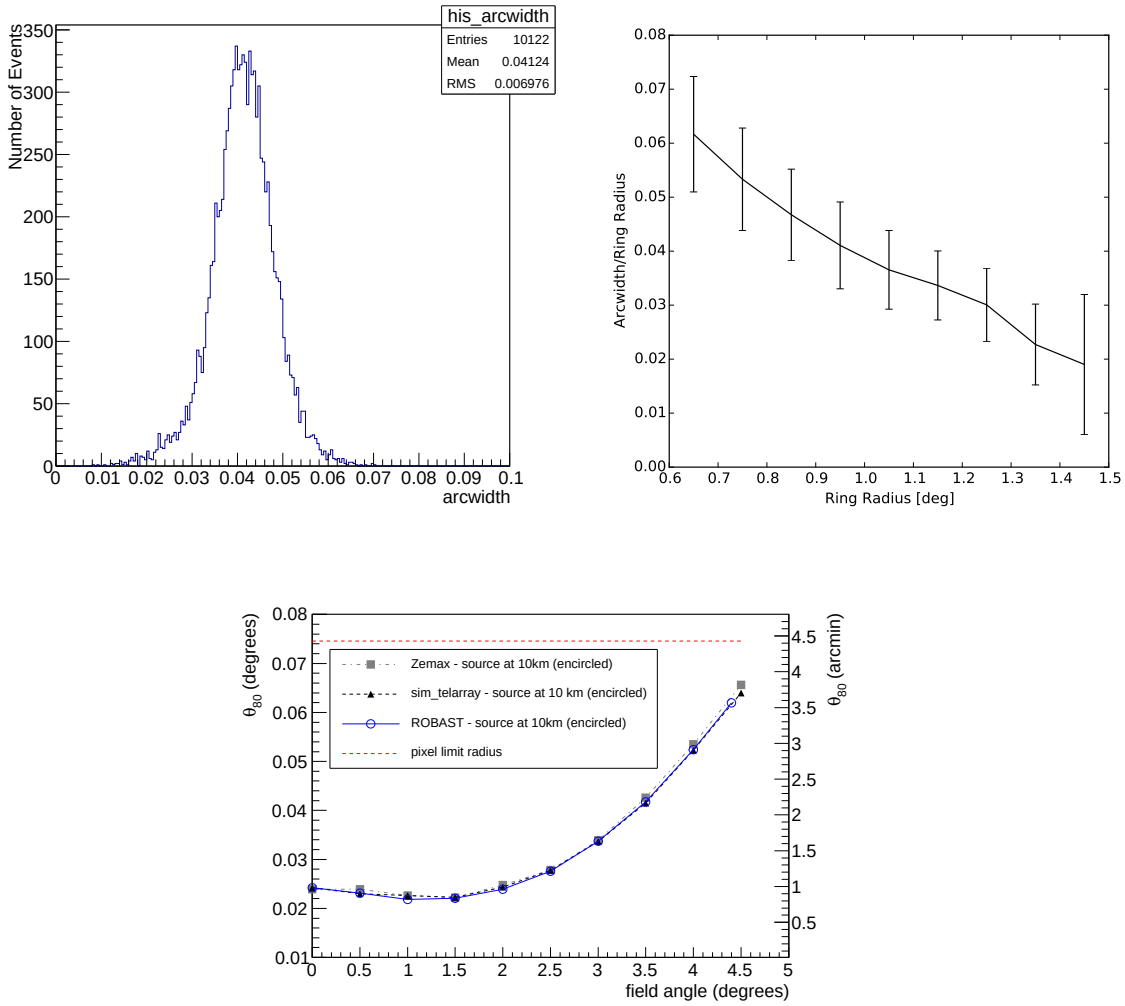


Figure 5.10: Top left: Distribution of the mean *Arcwidth* in each image for GCTS. Top right: *Arcwidth* compared to the reconstructed radius, analogue to Figure 5.11, where for larger rings (and therefore higher energies) the main effect comes from mirror aberrations. Bottom: PSF derived from ray-tracing for GCT from [25].

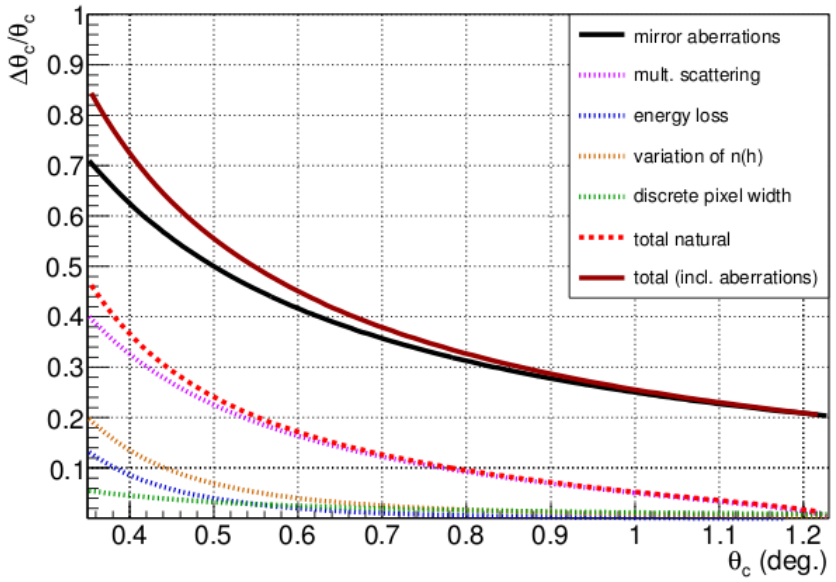


Figure 5.11: The individual components that contribute towards the blurring of the muon ring as calculated for GCT in [69]

components that lead to the ring broadening effect can be seen. It is clear that their effects are dependent on the ring radius (and therefore the energy of the muon). In Figure 5.10 the *Arcwidth* has been binned for different θ_C values, and for reference the PSF derived from ray tracing of starlight has also been included [25]. It can be seen that for larger muon radii, the *Arcwidth* approaches the expected on-axis PSF of GCT, which Figure 5.11 shows should be expected as the mirror aberrations become the dominant factor in the ring broadening.

5.8.6 Reconstructing Impact from Modulation

The final step required in the modelling process once again considers the distribution of light in the muon ring, this time concerning the modulation of light. Recall Figure 5.2, in which it was shown that the geometry of the relativistic muon track results in an asymmetric amount of light within the ring as a result of the different path lengths integrated over the primary mirror. By correctly modelling this light profile the impact position of the muon and therefore the muon efficiency can be reconstructed.

Returning to equation 5.6.16 in which the number of observed photons along

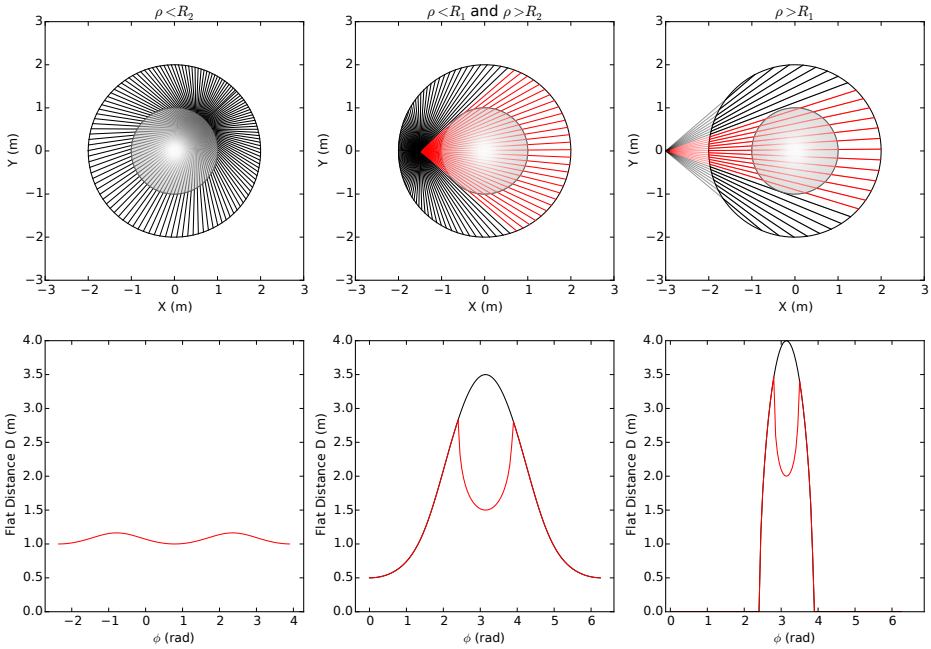


Figure 5.12: Modelling the modulation of light along the muon ring for three different cases: Left, muon lands with an impact distance less than the radius of the secondary; Middle, the muon lands with an impact distance between the radius of the secondary and the primary; Right, muon lands with an impact parameter greater than that of the primary. Top row shows each radial path crossing the mirror with the bottom row showing the resulting modulation around the ring. Red indicates an effect from shadowing of the secondary mirror.

the the ring, $dN_{obs}/d\phi$ is described as a function of ρ, the impact distance, and the angle with which the ring is traversed, ϕ. For now consider only the element that concerns the impact distance directly, $D(\rho, \phi)$, which for a single mirror system can be expressed as

$$D(\rho, \phi - \phi_o) = \begin{cases} 2R\sqrt{1 - (\frac{\rho}{R})^2 \sin^2(\phi - \phi_o)} & \text{for: } \rho > R \\ R[\sqrt{1 - (\frac{\rho}{R})^2 \sin^2(\phi - \phi_o)} + \frac{\rho}{R} \cos(\phi - \phi_o)] & \text{for: } \rho \leqslant R \end{cases} \quad (5.8.23)$$

where R is the radius of the primary mirror. However, in two mirror systems such as GCT, the effect of shadowing by the secondary mirror, the radius of which shall be defined as R', needs to be accounted for. For simplicity, consider only the case of on-axis muons; the effect of larger inclination angles acts to shift the position of the muon ring on the focal plane, therefore changing the region of shadowing. For small inclination angles the shadowing is still contained within the primary mirror as long as

$$\Delta m \sin(i) \; < \; R - R', \quad (5.8.24)$$

where Δm is the distance between the secondary and the primary mirror and i is the inclination of the muon. With the defined viewcone it is not expected that this will be exceeded. For the shadowing component, the definition of the path $D(\rho, \phi)$ as the chord that traverses the mirror unaffected by shadowing is retained and the term $D'(\rho, \phi)$ for areas that are affected is introduced, leading to a total effective distance that participates in focusing light as $C = D - D'$. An important parameter that will help distinguish between different cases is d, the perpendicular distance from the centre of the mirror to the line D, which is defined as:

$$d = \rho \sin(\phi). \quad (5.8.25)$$

In order to account for the secondary mirror, the three general conditions $\rho \; < \; R'$, $R' < \rho \leq R$ and $\rho < R$ must now be considered.

Impact within secondary

Recall in the case of a single mirror system, with $\rho \leqslant R$, the expression for the path length was given as

$$D(\rho,\phi) = R[\sqrt{1-(\rho/R)^2\sin^2(\phi)} + (\rho/R)\cos(\phi)] \tag{5.8.26}$$

which now gives the unobstructed component for the path distance. From this it is easy to conclude that the shadowing component should be similar, but instead substituting the radius of the secondary, R':

$$D'(\rho,\phi) = R'[\sqrt{1-(\rho/R')^2\sin^2(\phi)} + (\rho/R')\cos(\phi)] \tag{5.8.27}$$

where the total distance is $C = D - D'$. An example can be seen in the two left panels of Figure 5.12.

Impact between secondary and primary

If the muon now falls within the range $R' < \rho \leq R$ the term for $D(\rho,\phi)$ remains the same as in equation 5.8.26.

For the shadowing component there are now two possible situations. Firstly if the path does not traverse the secondary then only $C = D$ need be considered, which happens if the condition $d > R'$ is satisfied. In the alternative case, where the path is once again affected by shadowing, $D'(\rho,\phi)$ will be the same as the $\rho > R$ case in the single mirror system,

$$D'(\rho,\phi) = 2R'\sqrt{1-(\rho/R')^2\sin^2(\phi)} \tag{5.8.28}$$

Where again the total path length is set as $C = D - D'$ which can be expressed more generally as

$$C = \begin{cases} R[\sqrt{1-(\rho/R)^2\sin^2(\phi)} + (\rho/R)\cos(\phi)] & \text{for: } d > R' \\ R[\sqrt{1-(\rho/R)^2\sin^2(\phi)} + (\rho/R)\cos(\phi)] - 2R'\sqrt{1-(\rho/R')^2\sin^2(\phi)} & \text{for: } d \leq R' \end{cases} \tag{5.8.29}$$

This case can be seen in the middle two panels of Figure 5.12 where the paths affected by shadowing are marked in red.

Impact outside of primary

In the final case, where the muon track misses both mirrors, there now exist three situations depending on d. If the path does not traverse the secondary then the situation is the same as with the one mirror system. If instead the path intersects with the secondary then the same case is obtained for D' as in equation 5.8.28. Lastly, if the path does not cross the mirror at all then $C = 0$. This can be expressed as,

$$C = \begin{cases} 2R\sqrt{1-(\rho/R)^2\sin^2(\phi)} & \text{for: } d > R' \text{ and } d < R \\ 2R\sqrt{1-(\rho/R)^2\sin^2(\phi)} - 2R'\sqrt{1-(\rho/R')^2\sin^2(\phi)} & \text{for: } d \leq R' \\ 0 & \text{for: } d > R \end{cases} \tag{5.8.30}$$

where again an example can be seen in the right two panels of Figure 5.12.

Applying Models to Data

The corrected models for a two mirror system now need to be applied to the data in order to retrieve the impact parameter ρ along with the efficiency ϵ. To achieve this all pixels that are $\pm 0.5 \times D_{pix}$ of the muon ring are retrieved. These data are then fitted with equation 5.6.16 including a scaling factor to normalise to the individual pixel level such that the result can be compared to the observed values:

$$\frac{dN_{obs}}{d\phi}(\rho, \phi_o) = \frac{\alpha}{2} \cdot \frac{\omega}{\theta_c} \cdot \sin(2\theta_c) \cdot D(\rho, \phi - \phi_o) \cdot I \cdot T \cdot \epsilon_\mu \tag{5.8.31}$$

where ω is the pixel field of view. In the fit, the parameters ϕ_o, ρ and ϵ_μ are left free. An example of the fit can be seen in Figure 5.13. The resulting reconstructed impact parameter resolution can be seen in the left panel of Figure 5.14. From this it is clear that the current method of cut selection is not optimal. For events that land within the radius of the secondary mirror there seems to be an obvious difficulty in

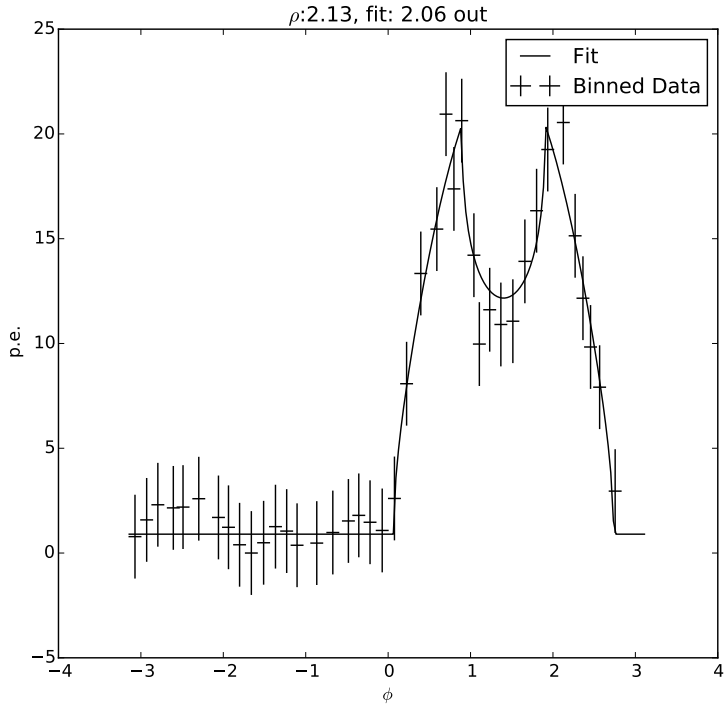


Figure 5.13: Fit of 5.8.31 to the binned pixels within ± 0.5° of the ring. Here the fit returns an impact distance of 2.06 m compared to the simulated distance of 2.13 m.

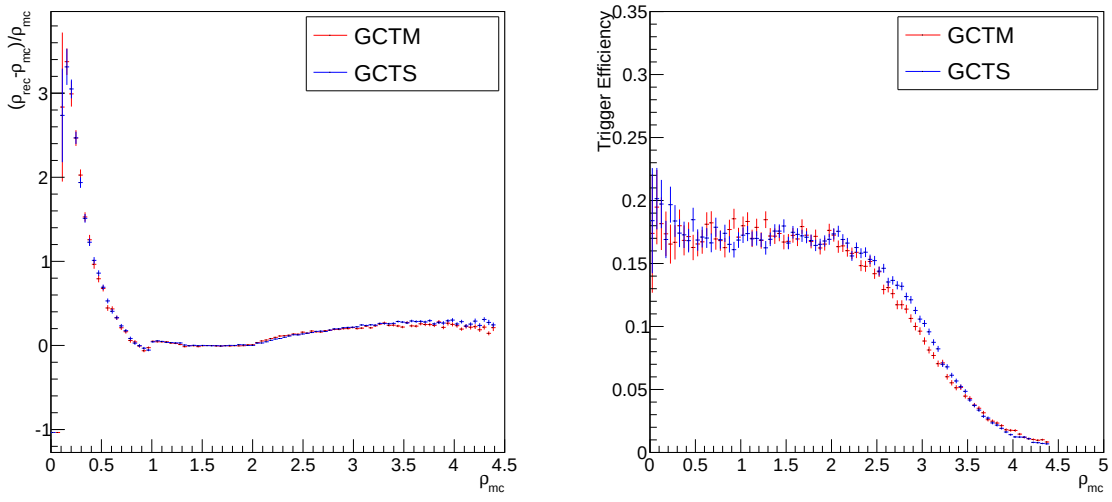


Figure 5.14: Left: Impact distance resolution as a function of simulated impact distance. Right: Trigger efficiency as a function of simulated impact distance where a severe drop past the 2 m radius of the primary can be seen. These include the edge cut, number of pixels >10, number of p.e. >40, and radius cut

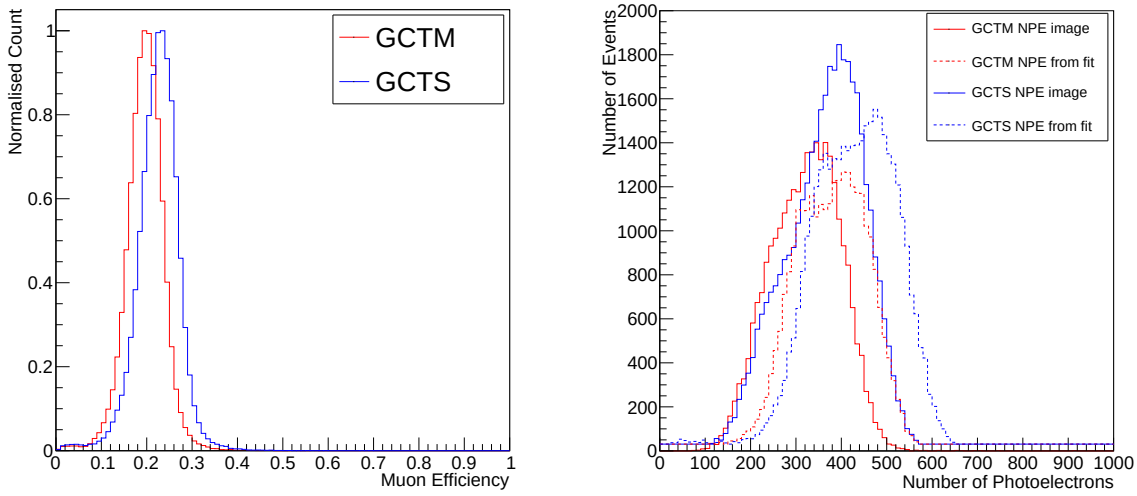


Figure 5.15: Left: Distribution of the efficiency, ϵ, along with the theoretical gamma efficiency calculated from the telescope configuration. Right: Distribution of the total number of p.e. from pixels retained after the tail-cut and the total returned from equation 5.6.14.

reconstructing the impact parameter. For events landing outside the primary mirror there is an increasing bias. In the right panel of 5.14, it can be seen that there is a flat trigger efficiency for events with an impact parameter less than the radius of the primary which tails off for events outside. It is therefore desirable to include a cut that rejects events that land outside of the primary mirror.

Continuing with the current method and cut selection, but now including a cut on impact distance of $\rho < R$, the obtained muon efficiency is now considered. In Figure 5.15 the distribution for the derived efficiency and the predicted number of p.e. for both GCTM and GCTS can be observed. From this it is determined that the efficiency, calculated from the muon reconstruction, of GCTM is (20.49±3.25)% and (23.50±3.50)% for GCTS (where the errors are one standard deviation of a Gaussian fit to the distribution). Using the inputs for the configuration files used in the simulations, it is possible to determine a theoretical efficiency against which this can be compared. For this, the program TESTEFF, available with the CORSIKA_SIMTELARRAY package, is used. Effectively it combines all the efficiency parameters such that

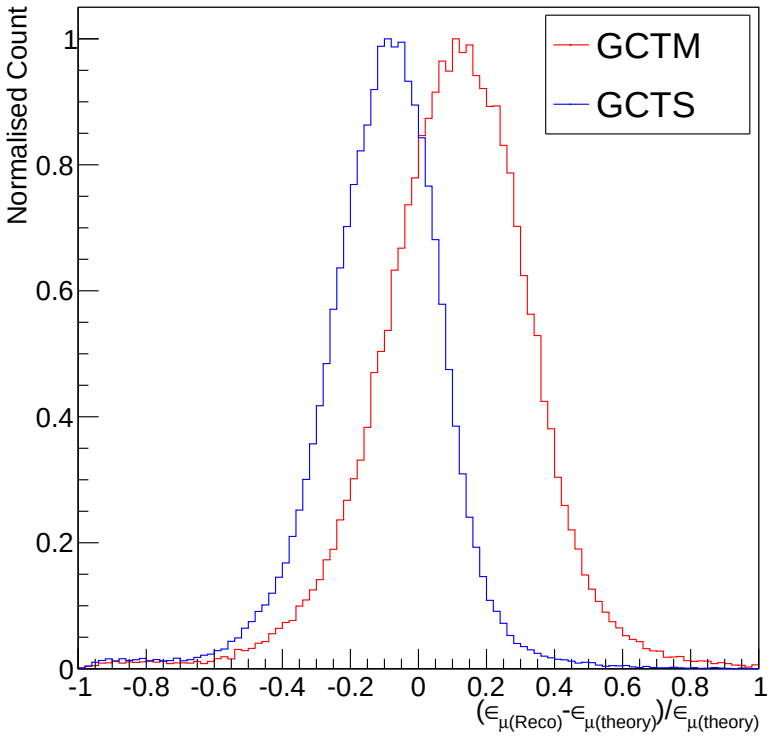


Figure 5.16: Expected rate of usable muons (passing all cuts) for GCTS, calculated using equation 5.9.33.

$$\epsilon_{\text{theory}} = T \cdot \epsilon_{\text{shadow}} \cdot \epsilon_{\text{mirr}} \cdot \epsilon_{\text{acceptance}} \cdot \epsilon_{\text{PDE}} \tag{5.8.32}$$

where the efficiencies are as described in Section 4.6 and have been evaluated here between 300 and 600 nm. From this it was determined that the theoretical efficiencies for GCTM and GCTS are 17.68% and 25.06% respectively and while these agree within one standard deviation of the reconstructed efficiency it seems that there is some bias that is introduced by the current reconstruction algorithm. This can be seen better by considering the distribution of the metric $(\epsilon_{\mu(reco)} - \epsilon_{\mu(theory)})/\epsilon_{\mu(theory)}$ which can be seen in Figure 5.16. Whether or not the bias is inherent to the telescope system or the fitting algorithm is unclear and will be investigated in future work. However, as long as its level is understood, it can be taken into account. The level of the bias is seen as an overestimate in the case of GCTM by 15.75% and an underestimate of 6.35% for GCTS.

	GCTM	GCTS
$\text{Rate}_{\text{good}}$ [Hz]	0.42	0.42
t_{obs} [min]	15.99	15.98
ϵ_{μ}	0.21	0.24
ϵ_{RMS} [%]	18.27	13.91
ϵ_{Bias} [%]	+15.75	-6.35

Table 5.1: Results obtained from Muon Analysis for events passing all cuts. Placing a cut on impact distance, such that events landing within the primary mirror, results in only complete rings being included.

5.9 Expected Rates and Summary Results

As a final part of the muon analysis it is important to know how long it will take to observe enough usable muon events (those passing cuts) with which a reliable calibration of the system can be performed. Therefore, using knowledge of the muon spectrum from Equations 5.4.6 and 5.4.9 and with the trigger efficiency shown in Figure 4.17 the expected rate can be calculated, using:

$$ER_{good} = \int_0^{\infty} F(E,h) \cdot R_{\text{imp}}^2 \cdot \pi \cdot 2\pi \cdot (1 - \cos(\theta_{\text{max}})) \cdot \eta(E) dE, \tag{5.9.33}$$

where $F(E,h)$ is the muon flux, R_{imp} is the maximum simulated impact radius, θ_{max} is the maximum simulated zenith angle (viewcone) and $\eta(E)$ is the overall trigger efficiency. Applying this to the data, an expected rate of $\sim$0.42 Hz is found for GCTM and GCTS. An initial estimate for the number of events required in order to derive the efficiency parameters was presented in [69] based on [36]. For an SST this was assumed to be 400 muon rings in order to obtain a RMS error on the efficiency of less than 20%, which corresponds to an observation time of around 16 minutes, easily achievable within a night. A summary of the final results can be seen in Table 5.1. While the results suggest muon calibration will be possible for both GCT configurations, it is only applicable if the method holds in the case of a deteriorated system, which will be the focus of the final section.

5.10 Monitoring System Degradation

One of the main goals of using muons as a form of calibration is to monitor the efficiency of the telescope system. Here it is investigated whether the method presented in this Chapter can still be used as the efficiency deteriorates. This loss in efficiency is expected to occur in multiple telescope elements over time, but, for simplicity this work has assumed that they can be combined and expressed as a loss of reflectivity in the mirrors[8]. In order to perform this evaluation, a similar data set to the one used in the rest of this Chapter was produced with CORSIKA. However, considering the drop in trigger efficiency for events falling outside the radius of the primary (recall Figure 5.14) and assuming those that do trigger can be rejected with a cut on reconstructed impact distance, only events landing within the primary mirror have been simulated. This increases the proportion of the data that is usable. The CORSIKA showers were then parsed to 10 instances of SIM_TELARRAY, each with an incremental increase in optical efficiency (from 1.0 to 0.1). For each of the data sets produced for both telescope configurations, the method described in Section 5.8 was applied.

In Figure 5.17, the resulting selection efficiency can be seen. For each configuration and for each level of deterioration it can be seen that, apart from at the lowest energies, there is no strong energy dependence to the selection efficiency. For both GCTM and GCTS, this also stays relatively constant up to around a 60% system optical efficiency. This consistency is important as, in order to perform calibrations from muons, the same population of muons must be sampled. Carrying this forward and using the selection efficiencies, the expected rate of good events can be evaluated for each level of mirror degradation, the result of which can be seen in the left panel of Figure 5.18. Compared to the results obtained earlier, the rate achieved for a fully operational system here is slightly larger at $\sim$0.47 Hz. This arises from considering only those muons landing within the primary radius, whereas in the former case

[8]Unless a more durable coating has been applied to the mirrors, it is assumed that this will comprise the major source of deterioration within the system. This is applied in the simulations with use of the SIM_TELARRAY parameter MIRROR_DEGRADED_REFLECTION

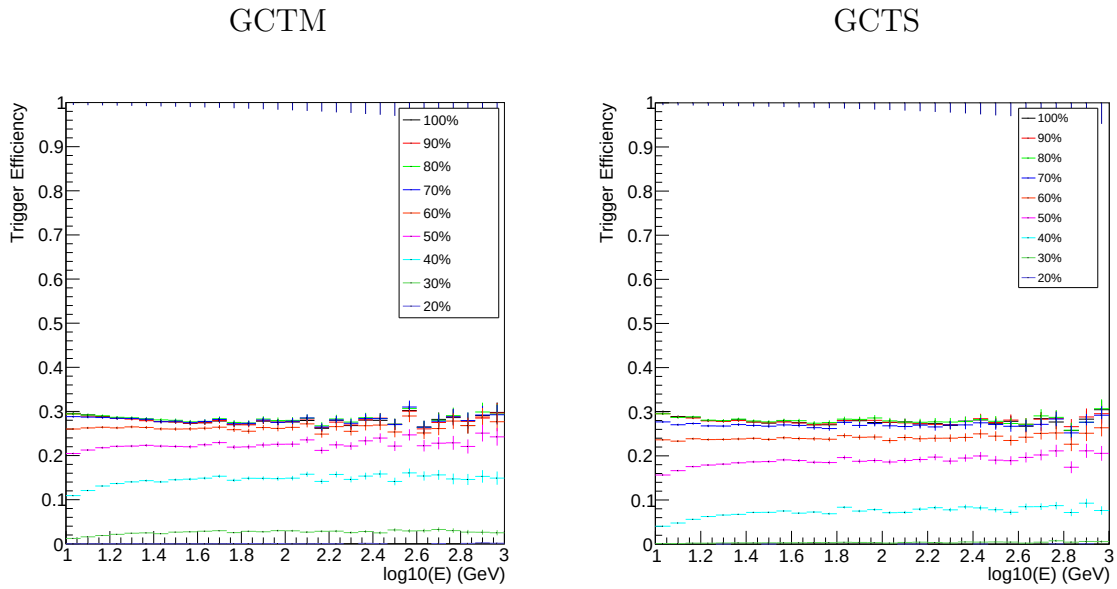


Figure 5.17: Selection efficiency of events passing all cuts for each level of degradation and each configuration.

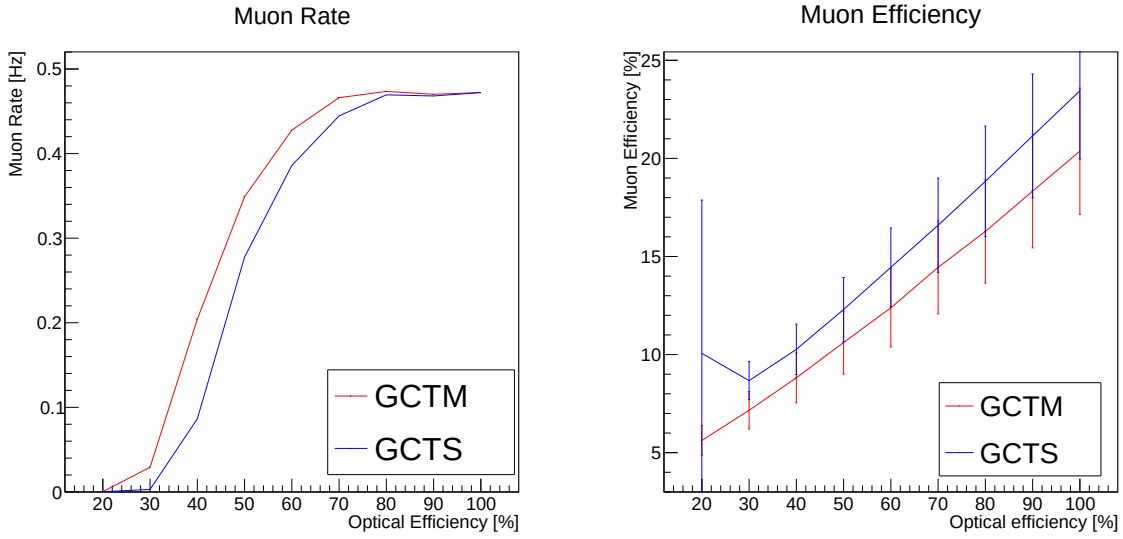


Figure 5.18: Left: The detection rate of usable muons for GCTM and GCTS as a function of optical efficiency. Right: The derived muon efficiency for both configurations as a function of optical efficiency.

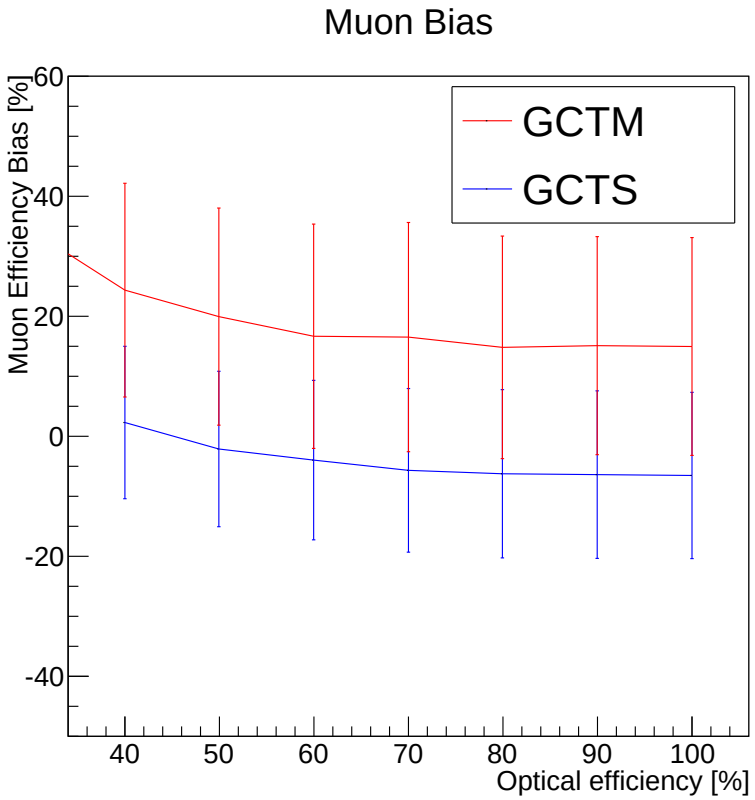


Figure 5.19: Derived muon efficiency bias $(\epsilon_{\mu(reco)} - \epsilon_{\mu(theory)})/\epsilon_{\mu(theory)}$ for GCTM and GCTS.

there will have been events which were reconstructed with a larger impact distance and therefore were rejected. With an improved algorithm for reconstructing the muon rings it can be expected that rates up to 0.47 Hz may be achievable.

In the right panel of Figure 5.18 the derived muon efficiency for a given optical efficiency is shown. There is a clear linear trend for both GCTM and GCTS. Considering the expected usable muon rate, it can be seen in Figure 5.18 that for GCTM the rate stays relatively constant down to an optical efficiency of 70% while GCTS is consistent only at optical efficiencies above 80%. The importance of this can be seen in Figure 5.19; while the RMS error at each degradation level remains below 20%, the level of bias increases as the telescope degrades. Further investigation and discussion within the CTA muon working group is required in order to arrive at an acceptable level of bias present in the estimation of the muon efficiency.

5.11 Conclusion

In this Chapter, a feasibility study looking at the possibility of using local unaccompanied muons as a form of calibration for GCT has been presented. Initially the method by which the important parameters of the muon shower are reconstructed, namely the energy from the ring radius, the shower inclination angle and direction from the position in the camera and finally the impact distance from the modulation of light around the ring, was presented. From this the selection efficiency of triggered muons passing several cuts was calculated, which in combination with knowledge about the muon spectrum, led to the determination of the expected muon rate. This was found to be very similar for both the telescope configurations, at 0.42 Hz. Assuming a minimum number of events required to obtain a statistical representation of the muon population of 400, this leads to an observation time of $\sim$16 minutes.

In this work, the observation strategy for observing muons has not been covered. As noted in Section 5.7.1, if IACTs operate in stereo mode, requiring that two or more telescopes trigger on a single event, muons would not be recorded. One solution for this is to carry out mono runs occasionally in order to record a sample of muons for calibration. However, for CTA to maximise its potential, this is undesirable

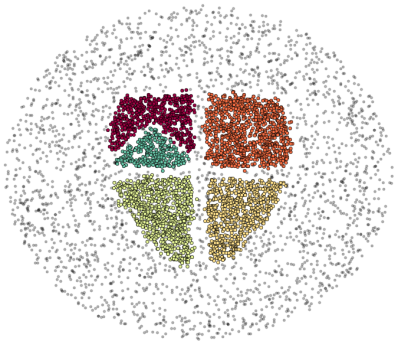
due to a loss in observation time and is not foreseen as an option, therefore a dedicated trigger needs to be included in the camera server which flags muon-like events. Studies within the muon working group suggest that it should be possible and future work will include an evaluation of a dedicated muon trigger scheme for GCT.

The main output from the muon calibration is the muon efficiency parameter which, along with the impact distance, is obtained from a fit to the modulation of light around the ring. While there is almost no bias present in the reconstruction in the radius of the muon ring, and a negligible bias present in the reconstructed impact parameter (at least when considering events landing within the primary and outside the secondary) there is a bias in the reconstructed efficiency compared to the theoretical efficiency (an overestimate by 15.75% for GCTM and an underestimate of 6.35% for GCTS). The source of the bias is not currently clear, however there are two possibilities. Firstly, the equations used for the modulation of light around the muon ring presented in Section 5.8.6 may not represent the most optimal solution of the two mirror design. The second possibility, due to the larger effect seen for GCTM, is that the bias introduced is an effect of the gaps between the camera detector modules. For a full understanding these will require further investigation.

The final work in this Chapter concerned the ability to reconstruct the muon efficiency parameter as the telescope system degraded, in this case assuming a loss in reflectivity of the telescope mirrors. In order to prevent selection bias of muon events, it is desirable to have a constant selection efficiency, and therefore rate, down to as low an optical efficiency as possible. In this preliminary work it was shown that rate stays constant down to 70% for GCTM and 80% for GCTS; below this there is an increasing bias in the reconstructed muon efficiency. As the current level of cuts provide a reasonable muon rate, it should be expected that a more detailed consideration and optimisation should lead to a more constant, albeit lower muon rate.

Chapter 6

Source Detection in Sparse Data Sets: Initial Study



This Chapter looks at the possibility of using the clustering algorithm DBSCAN to search for seed sources within the *Fermi* VHE (>100 GeV) extragalactic data. It can be seen as a feasibility study for the method as well as searching for new sources that could be followed up by ground based gamma-ray instruments. This also has the possibility to add to the science case for the CTA extragalactic key science project. The main body of this work has been published in Monthly Notices of the Royal Astronomical Society [24] and here some expansions are included to provide further detail.

6.1 Introduction

The *Fermi* space based gamma-ray observatory, since its launch in 2008, has been sweeping its Large Area Telescope (LAT) across the entire sky approximately once every 3 hours, slowly building up the most detailed view of the high energy sky to date. Every photon with an energy from sub MeV to around 300 GeV[1] that passes through the the LAT detector imparts arrival information within the Tracker, and deposits its energy within the Calorimeter (see Chapter 3.2 for a detailed description). The data recorded by *Fermi* are processed and then released to the scientific community as a large multi-dimensional database, which provides us with a wealth of information about the gamma-ray sky.

While the data are publicly available, the *Fermi* collaboration performs its own analysis, most notably releasing catalogues of significant sources. The *Fermi*-LAT 2 year point source catalog (2FGL), which was released in 2012 [9], was the main source of information for gamma-ray sources until the recent release of the 3rd source catalog (3FGL [10]) in 2015. This represents the deepest catalogue available for high energy sources, covering 4 years of observations over an energy range of 100 MeV to 300 GeV.

For each catalogue, the main input of potential sources was the source locations found previously by *Fermi* (including those which at the time did not pass the significance cut). For a potential new source, the detection technique involved the use of the wavelet analysis algorithms *mr_fitter* [113] and *PGWave* [50] to look for "seeds". These operate by looking for local deviations from the background model in the data returning a map of significant features. This is followed by a localisation step using tools such as pointlike [81] and pointfind [4] in order to obtain the most likely source location. Finally a likelihood fit is performed (see Section 6.3) in order to determine the significance and best spectral fit to the source.

In addition to the wavelet algorithms, a minimum spanning tree was also used on the extragalactic data above 10 GeV (the energy restriction was imposed as the

[1] Note that this was true for Pass 7 data. As will be seen in the following Chapter, the maximum energy now extends much higher.

algorithm does not account for "structured background"). However, in the last two decades and over many disciplines, there has been a substantial amount of work on clustering analysis as a major statistical technique for classifying large data sets into meaningful subsets. These methods are worthy of further investigation as potential source-finding algorithms for the LAT data set.

Alongside the minimum spanning tree clustering performed in the source detection for the 2FGL, investigation into clustering performance for *Fermi* was carried out in [116] using the density-based clustering algorithm DBSCAN (Density-Based Spatial Clustering of Applications with Noise [65]). By applying the cluster analysis to simulated *Fermi*-LAT data, the work done in [116] was able to show the statistical robustness of the code's ability to identify potential sources in noisy regions. The method has also been used to search for evidence of the 130 GeV emission towards the Galactic centre [47]. Therefore, in this Chapter, an initial feasibility study using the algorithm DBSCAN on real *Fermi*-LAT data will be presented.

In order to accomplish this, a choice was made to apply the cluster analysis to all $E_\gamma \geqslant 100$ GeV photons with a Galactic latitude, b, that exudes the Galactic plane ($|b| > 10^\circ$). Firstly, since the extragalactic diffuse background has a spectral index of 2.41, complications due to background noise, which mainly affect lower energies, will be reduced [5]. Secondly, as the computational complexity of DBSCAN runs as $O(n^2)$, by using only the high-energy events a full, unbiased and model-independent clustering analysis of the whole sky can be performed without using a large amount of computing time[2]. Finally, the possibility of increasing the known VHE γ-ray population of TeVCat[3] active galactic nuclei (AGN) from its current number of 61 is attractive, particularly in the light of framing the scientific priorities for the forthcoming Cherenkov Telescope Array [12], [112].

This Chapter is organised as follows. In Section 6.2 an overview of the major clustering algorithm types will be given, leading to Section 6.2.1 in which the chosen algorithm for this work, DBSCAN, will be described together with its application to

[2]It is possible to improve the speed up to $O(n \log n)$ by pre-computing the *EPS*-neighbourhoods (see Section 6.2.1). However, the computational demand of this work did not require this step.

[3]Online catalogue of VHE ground based detections `http://tevcat.uchicago.edu/`

the *Fermi*-LAT VHE data. In Section 6.3 the maximum likelihood method used to verify sources and the sources found in this work will be shown. Section 6.4 will present a discussion of the performance of the DBSCAN algorithm and the sources it found with a conclusion in Section 6.5.

6.2 Clustering Algorithms

The term clustering algorithm is rather general and essentially encompasses a broad range of different algorithms covering a wide variety of uses. For simplicity they can be divided into 5 main categories:

Hierarchical

These types of clustering algorithms are best used in the classification of data, iteratively dividing data into groups using either a top-down or bottom-up method. For example, a bottom-up clustering algorithm would start with all data points as separate clusters, then each iteration would find similarities between these data sets and build subsequent levels of clusters, each iteration reducing the number of clusters until all data are combined into one lead cluster. The resulting data can then be expressed in a dendrogram and then a cut can be applied to obtain a number of sub categories [18].

Partitioning

In this class of clustering algorithms, data points are iteratively relocated between a predefined number of clusters. With each iteration, the centres of each cluster are recalculated based on the current members and then points are reassigned to the closest cluster. Often these work by minimising a metric known as the partitioning error (which can be defined in a variety of ways) and either complete after a set number of iterations or when the error reaches a minimum. While the algorithms are generally simple and fast, the requirement for a predefined number of clusters and the sensitivity to the initial positioning is a large downside. The most well known algorithm of this type is the K-means algorithm [71].

Model-based

Unlike other algorithms, these not only identify groups of clusters but also find characteristic descriptions for each group. These are often based on neural networks or decision trees and depend on weight vectors between data points. While sensitive to initial weight conditions, these algorithms are useful for multi dimensional data [127].

Grid-based

Algorithms that use a grid-based system work in a slightly non-traditional way. Instead of considering the data points individually, their properties are grouped on a grid which is then sorted and the cluster centres are determined. This can reduce the computational demand by simplifying the system [14].

Density-based

Finally, density-based algorithms produce clusters if the number of events within a certain area is greater than the number in its surroundings, either through density-connected points or based on an explicitly defined density function. One of the simplest and most widely-used examples of this is DBSCAN [65]. Its ability to pick out clusters of arbitrary shape from noisy data and its use in previous preliminary studies [116], [47] make it the logical choice for examining clustering in the *Fermi*-LAT data.

6.2.1 DBSCAN

The density-based algorithm DBSCAN was first described in [65] and was designed to combat the problem of efficiently detecting clusters of arbitrary shape in large noisy databases. The main aim of DBSCAN is to perform spatial analysis, by building an algorithm that depends on only one parameter, the density of the data. In practice, however, this necessitates the use of two input parameters, a search radius *EPS* around a given point within which a number of points, defined as *MinPts*, would be required to represent an over density and therefore the seed of a cluster.

From this, the central logic of DBSCAN is created, defining the idea of a core point and core samples. These are defined as a point p, which satisfies the condition $N_{EPS}(p) \geqslant MinPts$. That is, p is a core point if the number of events within its *EPS*-Neighbourhood is equal to or greater than that given by the *MinPts* parameter.

The work in this chapter uses code built on the readily available Scikit-Learn PYTHON library [101] in which the clusters are computed as follows:

1. for each point p in a set of objects D, the number of points within the *EPS*-Neighbourhood ($N_{EPS}(p)$) is found;

2. if the core sample condition $N_{EPS}(p) \geqslant MinPts$ is satisfied, then p is a core point and is added to the cluster C;

3. if a point q within the *EPS*-Neighbourhood of the core point p also satisfies the core sample condition then p and q are *density-connected* and q is added to C. If not, it is classified as a *border point* or *density-reachable*;

4. step 3 is repeated for every candidate core point for C;

5. the algorithm moves to a new, unprocessed, core point and returns to step 2.

In brief, the DBSCAN algorithm takes every point, considers whether it is in a dense region and then builds up a cluster by adding all nearby points, with respect to *EPS* and *MinPts*, that exist above a certain threshold density. All objects that have been processed but are not considered density-connected to a cluster are defined as noise. This facilitates the creation of clusters of arbitrary shape, as the process of growing a cluster does not depend on the shape of the current protocluster, and it also allows the efficient rejection of background events. For a full example of this process, see Figure 6.1.

No clustering algorithm is without its drawbacks. One of the difficulties of using DBSCAN is that the initial choice of *EPS* and *MinPts* can strongly affect the outcome of the algorithm. However, a statistical analysis performed on simulated *Fermi* data presented in [116], showed that an optimum choice of input parameters was achievable for this application. At least for the *EPS*, it was concluded that a sensible

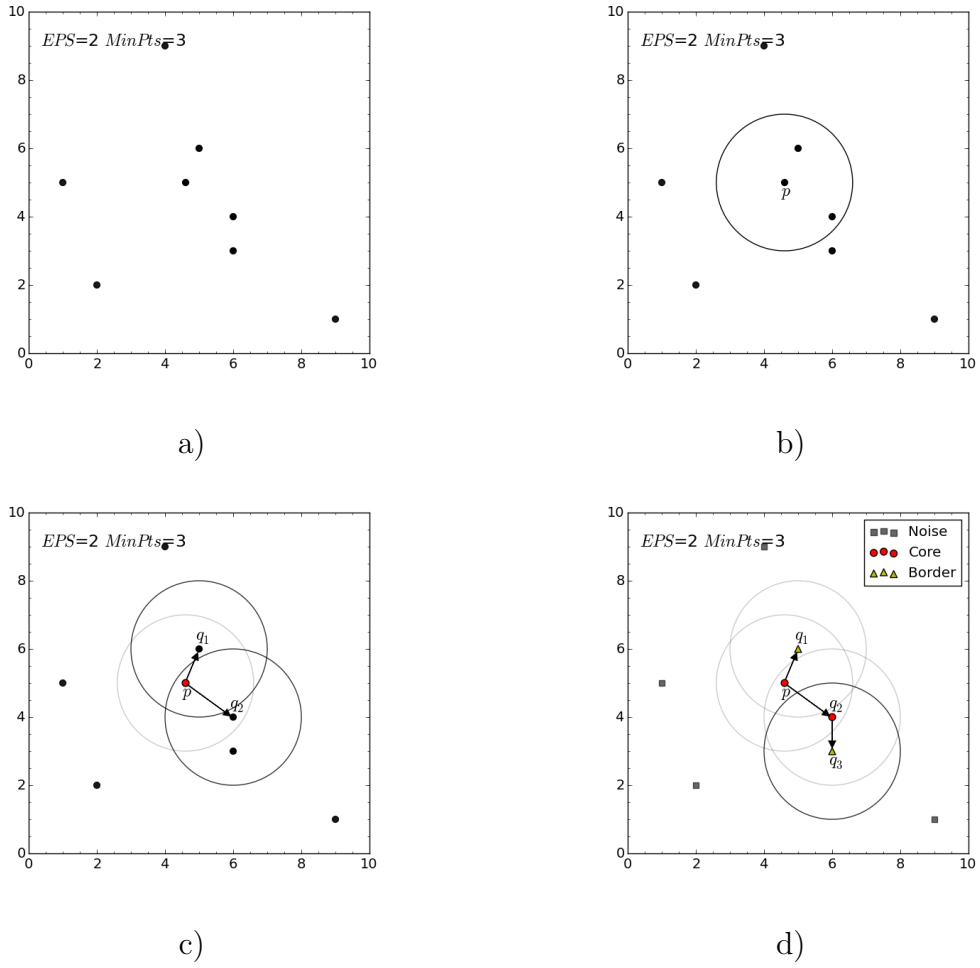


Figure 6.1: Description of DBSCAN with EPS=2 $MinPts$=3. Starting with point p, which is classified as a core point as there is a total of three points within the EPS neighbourhood, q_1, q_2 and p itself. q_1 is directly density-reachable from p but does not itself satisfy the core condition and is therefore defined as a border point. q_2 is also directly density-reachable from p but is a core point. The classification of q_3 follows that of q_1. All the rest of the points are considered as noise. The resulting cluster has 2 core points and a total of 4 points, including border points.

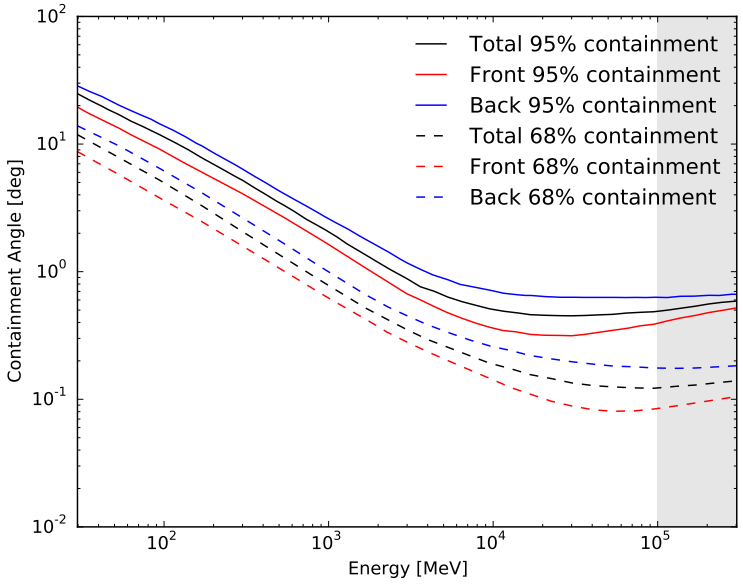


Figure 6.2: The point spread function for the *Fermi* LAT Pass 7 instrument response function, showing the 68% and 95% containment radii for FRONT and BACK converting events along with the total. The shaded area represents the energy range used in this work. Adapted from https://www.slac.stanford.edu

choice would be the point spread function of the *Fermi*-LAT detector, seen in Figure 6.2. In the case of this work, for clustering above 100 GeV, as the *Fermi*-LAT response functions give a PSF of 0.12° at 100 GeV for a 68% containment radius and 0.5° for 95% [8] (for both FRONT and BACK converting events, see Figure 6.2), a range between these values was chosen.

A second limitation of DBSCAN is its inability to deal with a spatially non-uniform background. In these cases, the intrinsic cluster structure may be masked by a non-ideal global set of parameters. For example, the algorithm may fail to reveal substructure in areas of general over density in favour of finding clusters in fainter regions. Conversely, it may sacrifice these more diffuse clusters in order to obtain a characterisation of the bright area. There are clustering algorithms available (e.g [20] and [104]) that modify DBSCAN to allow for its application to data with variable noise. Alternatively one could run DBSCAN in a scanning mode, adjusting the input parameters for each scan region, which was the approach taken in both [116] and [47]. For our application of clustering off-plane at energies greater

than 100 GeV, the variation in the diffuse background is greatly reduced, to the point where it can be considered negligible and for now this is assumed to be the case.

6.2.2 Clustering of VHE Gamma-Ray Events

The VHE domain provides a good testbed for the validation of DBSCAN. By restricting ourselves to energies greater than 100 GeV, this not only reduces the problem of varying background, but also the computational power needed to perform an unbiased clustering search of the whole sky. With its long exposure time and full sky coverage, *Fermi* gives us access to the deepest extragalactic scan presently available at these energies. Indeed, recent work took advantage of *Fermi*-LAT's deep exposure to discover two new VHE-bright AGN [43,44]. It is important to note, however, that these studies only searched for VHE emission around bright, spectrally hard, *Fermi*-LAT detected BL Lac objects. Given the relatively small number of known VHE gamma-ray objects, it is important that the statistical methods should be investigated in the context of a model-independent search, which could lead to greater understanding of VHE populations.

For the data set, all *Fermi*-LAT events for the first 6.25 years of operation from 4th August 2008 to 28th November 2014 (Mission Elapsed Time: 239557417 to 438847466) were taken for events with energies greater than 100 GeV, including both FRONT and BACK converting SOURCE classes. The Galactic plane ($|b| < 10°$) was also excluded from the scan as the source confusion resulting from the poor angular resolution makes reliably picking out individual clusters in this dense region difficult.

In accordance with the Pass 7 rep criteria, a zenith cut of 100° was applied to the data to remove any γ-rays induced by cosmic ray interactions in the Earth's atmosphere. The good time intervals were generated by applying a filter expression of "`(data qual == 1) && (lat config == 1) && abs(rock angle)` <52°" to the data using the *Fermi* tool GTMKTIME, where the (data qual) and (lat config) flags remove sub-optimal data affected by spacecraft events and the (abs(rock angle)) flag removes data periods where the LAT detector rocking is greater than 52°. These

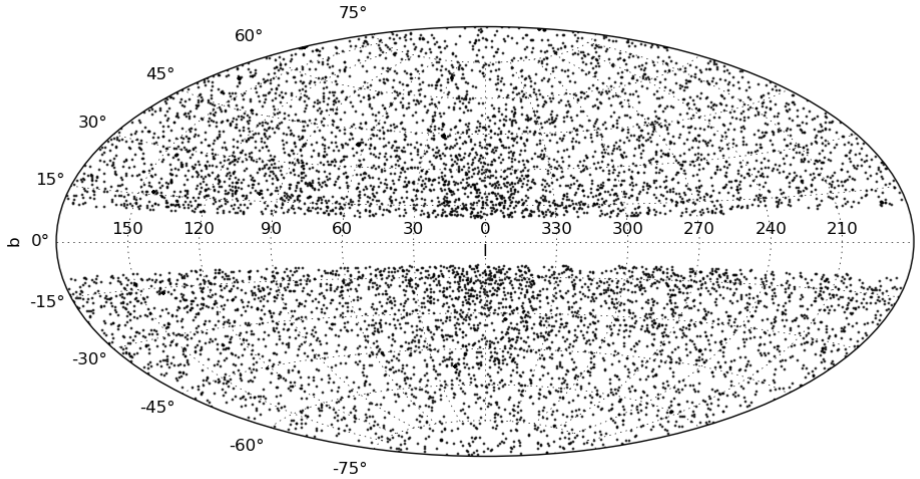


Figure 6.3: Distribution of events above 100 GeV passing all cleaning cuts and with the Galactic plane removed.

criteria are summarised in Table 6.1. The events remaining after cleaning on which the clustering analysis is performed can be seen in Figure 6.3.

Based on [116] a range of *EPS* values between the 68% and 95% containment radii was chosen for the clustering parameters. Using the the Pass 7 response files for the PSF at 100 GeV for both FRONT and BACK converting events, this was found to equate to a range of $\sim 0.12°$ to $0.5°$. Due to the relatively low statistics, *MinPts* was chosen to be the minimum number of events that could constitute a cluster statistically, namely 3 events.

For each cluster, the effective radius from the cluster centroid was calculated as $r_{eff} = \sqrt{\sigma_x^2 + \sigma_y^2}$, where σ_x and σ_y are the uncertainties expressed as the standard deviations in the event position. To determine the significance of the cluster the Likelihood Ratio Test (LRT) was applied, as described in [85] and applied in both [116] and [47],

$$s = \sqrt{2\left(N_s \ln\left[\frac{2N_s}{N_s + N_b}\right] + N_b \ln\left[\frac{2N_b}{N_s + N_b}\right]\right)}, \tag{6.2.1}$$

where N_s is the number of events taken from the DBSCAN and includes core and border events. The background N_b was estimated from the number of events between 2 r_{eff} and 3 r_{eff}. A cluster significance of $s = 2$ was set as the minimum

Cut Name	Value
tools version	v9r31p1
response function	Pass 7 rep.
Emin	100 GeV
Emax	300 GeV
T_{start} (MET)	239557417
T_{end} (MET)	438847466
zenith	100°
evclass	SOURCE (2)
conversion type	FRONT & BACK
DATA_QUAL	1
LAT_CONFIG	1
ABS(ROCK_ANGLE)	<52°

Table 6.1: Selection criteria for the data to which the clustering analysis was applied.

significance for a cluster. When N_s and N_b are large, which is not the case here, this represents a fluctuation of 2σ above the background. Therefore the LRT is used only as an indicator (the validity of this assumption is discussed in Section 6.4.1).

A study of the effects of changing *EPS*, which will also be described in Section 6.4.1, showed the optimal value to be 0.4°. Using this and a *MinPts* of 3, a total of 49 clusters were found with an LRT significance of $s > 2$ (See Figure 6.4) in 6.25 years of *Fermi*-LAT data above 100 GeV. Of these 21 are spatially coincident with existing TeVCat sources and 15 of the remaining 28 coincide with 3FGL sources. The TeVCat sources can be found in Table 6.2 and seen in Figure 6.5 for reference and will not be considered further in this work. In order to verify the potential VHE sources, it is necessary to utilise the *Fermi* analysis tools to perform a maximum likelihood analysis.

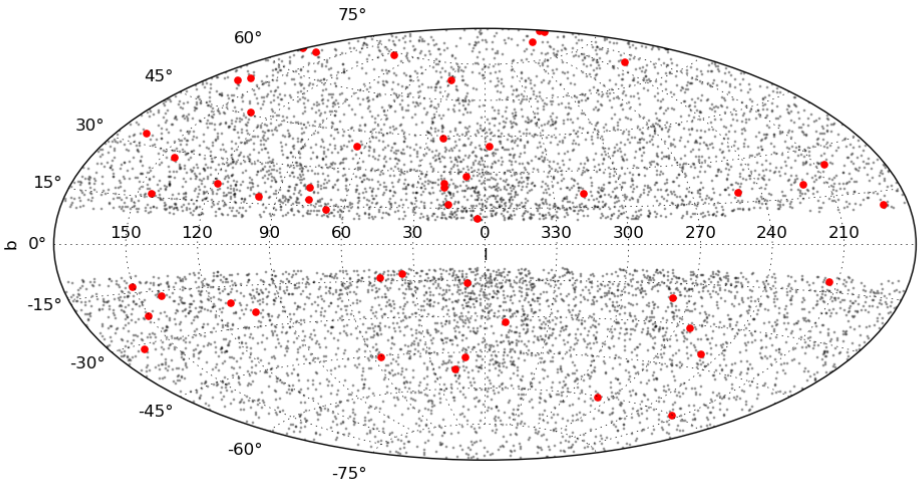


Figure 6.4: Clusters found using DBSCAN and their spatial position within the total data set.

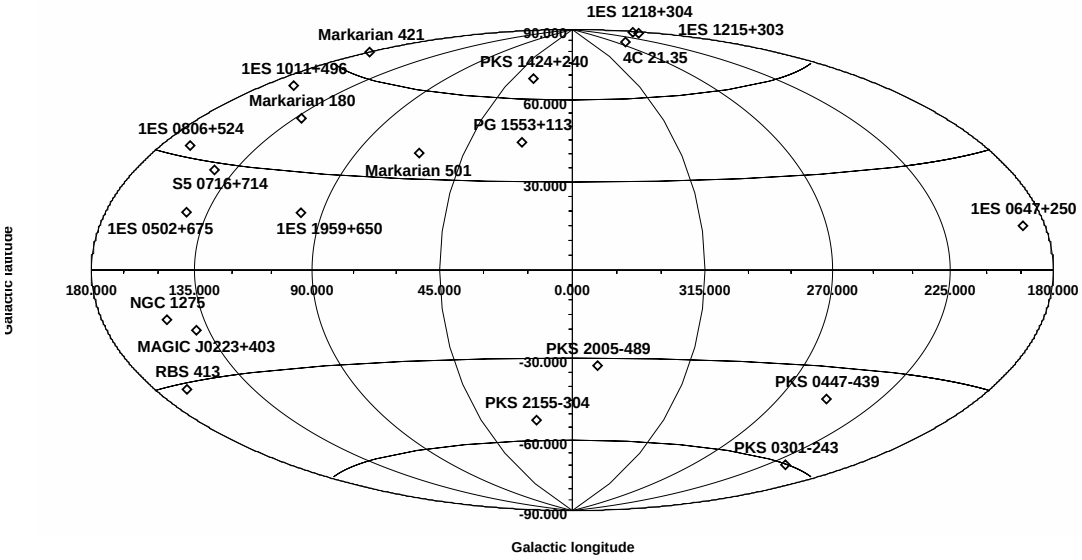


Figure 6.5: Clusters found that are spatially coincident with known VHE emitters that are within the TeVCat catalogue. These are shown for reference but are not analysed further.

	Fermi ID	Counterpart ID	$n_{0.4}$	$s_{0.4}$	TS 100-300 GeV	Flux 100-300 GeV $\times 10^{-11}$ ph cm^{-2}s^{-1}
1	3FGL J0222.6+4301	MAGIC J0223+403	11	3.91	133.68	6.11±1.77
2	3FGL J0303.4-2407	PKS 0301-243	8	3.33	76.80	4.20±1.59
3	3FGL J0319.8+1847	RBS 413	3	2.04	24.37	<2.13
4	3FGL J0319.8+4130	NGC 1275	3	2.04	24.75	<2.06
5	3FGL J0449.4-4350	PKS 0447-439	7	3.12	67.65	3.47±1.34
6	3FGL J0508.0+6736	1ES 0502+675	13	4.12	161.72	5.75±1.50
7	3FGL J0650.7+2503	1ES 0647+250	8	3.33	64.54	4.01±1.54
8	3FGL J0721.9+7120	S5 0716+714	5	2.63	39.15	2.03±0.91
9	3FGL J0809.8+5218	1ES 0806+524	3	2.04	31.85	1.32±0.77
10	3FGL J1015.0+4925	1ES 1011+496	13	4.02	153.57	6.86±1.82
11	3FGL J1104.4+3812	Markarian 421	95	11.53	1259.24	50.0±4.97
12	3FGL J1136.6+7009	Markarian 180	5	2.63	31.05	1.14±0.67
13	3FGL J1217.8+3007	1ES 1215+303	5	2.63	28.96	2.15±1.11
14	3FGL J1221.3+3010	1ES 1218+304	9	3.53	83.24	4.79±2.57
15	3FGL J1224.9+2122	4C 21.35	3	2.04	30.09	1.66±0.98
16	3FGL J1427.0+2347	PKS 1424+240	9	3.53	81.44	4.64±1.64
17	3FGL J1555.7+1111	PG 1553+113	27	6.08	287.47	14.3±2.83
18	3FGL J1653.9+3945	Markarian 501	47	8.11	502.40	22.4±3.34
19	3FGL J2000.0+6509	1ES 1959+650	9	3.53	52.76	3.85±1.30
20	3FGL J2009.3-4849	PKS 2005-489	9	3.53	74.93	4.14±1.48
21	3FGL J2158.8-3013	PKS 2155-304	21	5.31	218.82	12.4±2.75

Table 6.2: 21 Sources found at E $\geqslant$ 100 GeV with DBSCAN which are also in the TeVCat and 3FGL catalogues. Here we show the 3FGL and TeVCat identifiers, the number of events found with DBSCAN, the LRT significance returned and the Test Statistic (TS, described in following section) and flux from the likelihood fit.

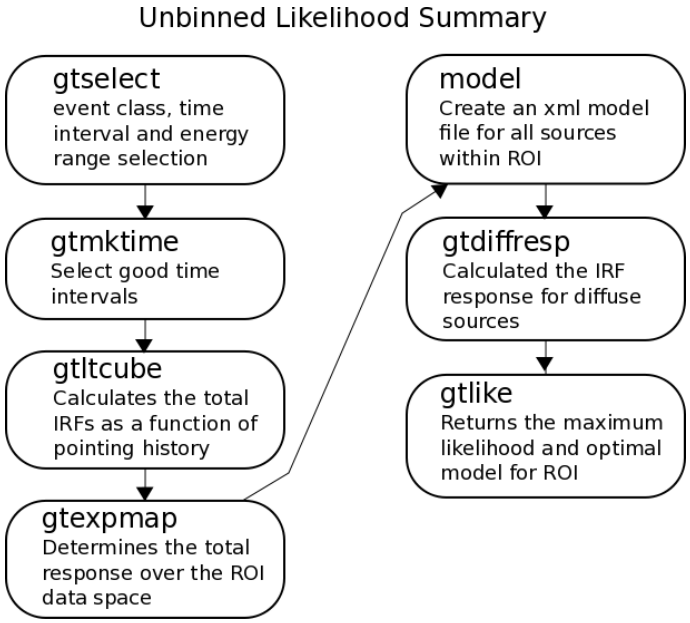


Figure 6.6: Simplified summary of the *Fermi* unbinned annalysis chain

6.3 Verification of VHE Clusters using *Fermi* Analysis

For each significant cluster found using the DBSCAN algorithm, the full 6.25 years worth of the *Fermi*-LAT data within a ROI of radius 5° surrounding the cluster position was used for further analysis. As before, the data were reduced with the *Fermi* tools GTSELECT and GTMKTIME in order to apply a zenith cut and to keep only the "good time intervals" according to the same Pass 7 criteria for SOURCE class events between 100 and 300 GeV (see Table 6.1).

An unbinned likelihood analysis was run on each source (a summary of this process can be seen in Figure 6.6), modelling each cluster with a power law spectral shape of the form,

$$\frac{dN}{dE} = A \times \left(\frac{E}{E_o}\right)^{-\Gamma}, \tag{6.3.2}$$

where A is the normalisation, Γ the spectral index and E_o the scaling factor. In addition to modelling the cluster, each analysis used a model file consisting of all point sources within 15 degrees of the cluster position, as well as the corresponding Pass 7 Galactic and extragalactic diffuse models (gll_iem_v05_rev1.fit and iso_source_v05.txt respectively). The position and the spectral shape of these point

sources were taken from the 3FGL. Furthermore, several clustering events were found to be located in close proximity to known extended γ-ray sources, namely W28, W30, W44, the Cen A lobes and HESS J1841-055. These extended sources were accounted for with their respective spatial distribution models from the 3FGL. During the analysis, the normalisation and the spectral index of the cluster source and the point sources within the ROI where left free. Modelled sources outside the ROI but within 15° of the cluster had their parameters frozen to those published in the 3FGL[4]. Likewise, the normalisation factor of the extragalactic diffuse emission was left free, and the Galactic diffuse template was multiplied by a power law in energy, the normalisation of which was allowed to vary[5].

From the unbinned analysis with the above model, a best-fit power law model was derived and the integrated flux for each cluster was obtained along with the resulting likelihood Test Statistic (TS). The TS is an indication of the source significance, with a general assumption that $\sigma \approx \sqrt{\mathrm{TS}}$, which is defined as

$$\mathrm{TS} = -2 \ln(L_{max,0}/L_{max,1}), \tag{6.3.3}$$

where $L_{max,0}$ is the maximum likelihood value for a model without an additional source (the 'null hypothesis') and $L_{max,1}$ is the maximum likelihood value for a model with the additional source at a specified location. If the analysis returned an insignificant result (TS <25) for $E_\gamma \geqslant 100$ GeV, upper limits for the flux were calculated using the *Fermi*-LAT Python likelihood analysis package.

To ensure that there where no γ-ray sources in the ROI that were not taken into account in the model (such as other sources not included in the 3FGL), the *Fermi* tool GTTSMAP and the best fit model were used to create a TS significance. An example can be seen in Figure 6.7 and shows the TS value distribution within the ROI. Apart from the central source there is no other significant source within the field of view, with two exceptions: 1ES 1011+496, which is located 2.89° away from 3FGL J1031.2+5053, and Markarian 421 which is 5.08° away from 3FGL J1120.8+4212.

[4]In some cases extra parameters were frozen in order to improve the global fit. Sources with a significance less than 5 had their parameters frozen, sources with a TS <1 were removed altogether

[5]`http://fermi.gsfc.nasa.gov/ssc/data/access/lat/BackgroundModels.html`

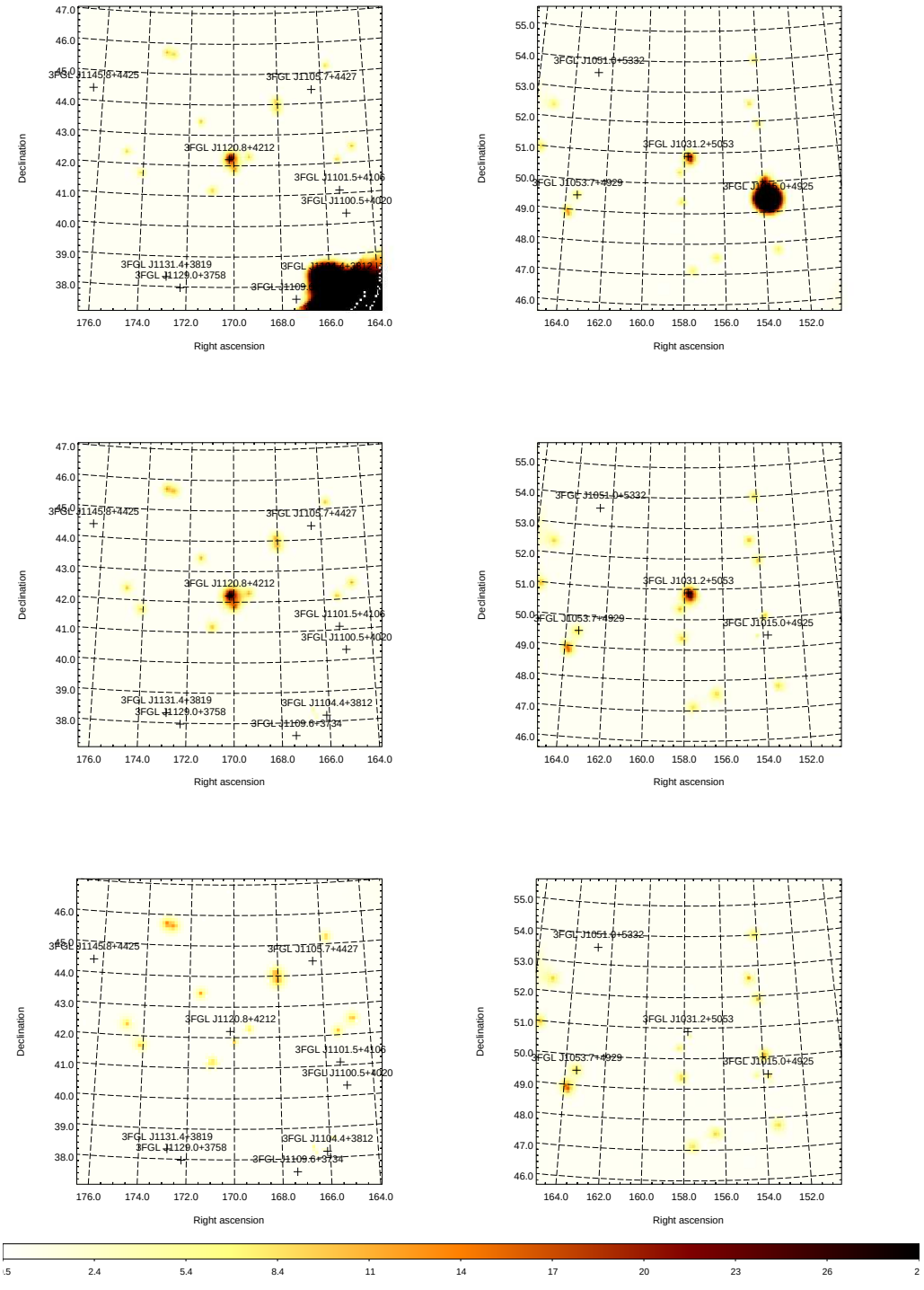


Figure 6.7: Example TS maps for 3FGL J1120.8+4212 and 3FGL 1031.2+5053. The top row does not include any point sources in the model, therefore revealing any areas that exist as an excess above the background. This includes the central sources and Markarian 421 (left) and 1ES 1011+496 (right). In the second row Markarian 421 and 1ES 1011+496 have been correctly modelled, and in last row all sources have been accounted for. Colour scale here is TS.

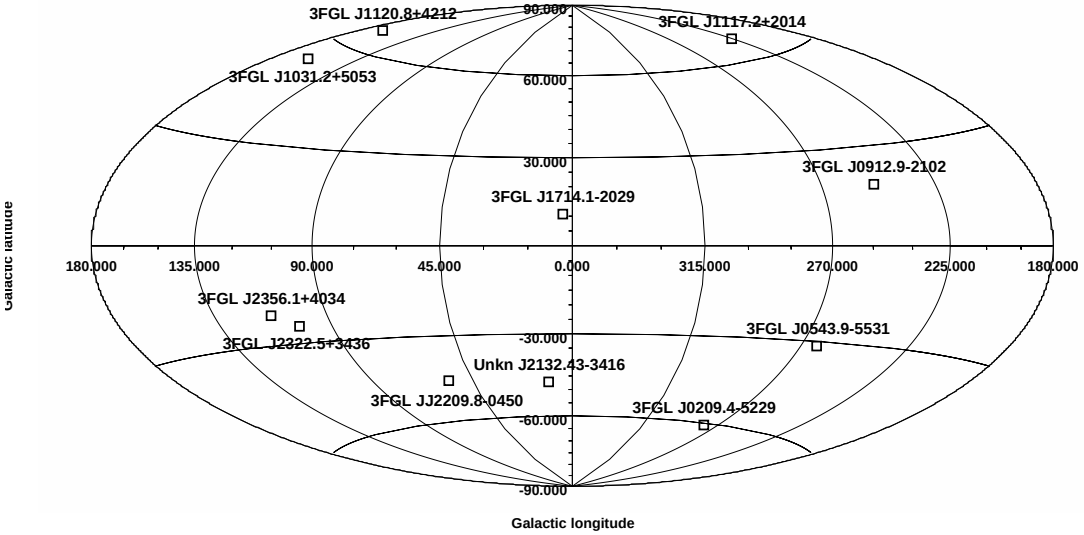


Figure 6.8: The 11 clusters that were found with a TS >25 above 100 GeV.

However, these have been accounted for and leave no residual in the fitted model.

Lastly, after accounting for all point sources within the field of view with the *Fermi* tool GTTSMAP, one final refinement of the model file was performed, i.e. the *Fermi* tool GTFINDSRC was used to determine a more precise localisation of the source's right ascension and declination. The differences between the GTFINDSRC results and the position found by DBSCAN all agree within the 95% PSF and in most cases to better than 0.1°. The resulting positions, fluxes and TS values of all 28 DBSCAN clusters can be found in Table 6.3.

6.4 Discussion

Using DBSCAN parameters $EPS = 0.4°$ and $MinPts = 3$ on 6.25 years of *Fermi*-LAT Pass 7 reprocessed data for $E_\gamma \geqslant 100$ GeV, excluding data from $|b| < 10°$, 49 sources have been found which return a significant likelihood ratio. Of the 61 extragalactic objects already existing in both the *Fermi*-LAT third point source catalogue (3FGL)

	Fermi ID	Counterpart ID	RA	Dec	z	TS	Flux $\times 10^{-11}$	TS	Flux $\times 10^{-9}$	Γ
						>100 GeV	>100 GeV	<100 GeV	<100 GeV	0.1-100 GeV
			deg	deg			ph cm^{-2}s^{-1}		ph cm^{-2}s^{-1}	
1	3FGL J0209.4-5229	RBS 285	32.45	-52.48	-	37.08	1.56 ± 0.74	690.92	7.56 ± 1.00	1.74 ± 0.053
2	3FGL J0543.9-5531	RBS 0679†	85.99	-55.55	0.273	51.12	2.07 ± 0.96	722.77	8.16 ± 1.06	1.72 ± 0.051
3	3FGL J0912.9-2104	MRC 0910-208	138.31	-21.09	0.198	36.09	2.34 ± 1.09	278.69	6.25 ± 1.47	1.83 ± 0.085
4	3FGL J1031.2+5053	RBS 877	157.74	50.88	0.360	27.97	1.59 ± 0.89	465.99	5.39 ± 0.030	1.77 ± 0.0024
5	3FGL J1117.0+2014	RBS 0958	169.24	20.25	0.138	36.21	1.94 ± 1.11	802.22	14.39 ± 0.38	1.95± 0.010
6	3FGL J1120.8+4212	RBS 0970†	170.16	42.26	0.390	34.34	2.18 ± 1.13	730.57	4.31 ± 0.53	1.55 ± 0.050
7	3FGL J2322.5+3436	TXS 2320+343	350.63	34.60	0.098	41.82	2.13 ± 1.08	76.50	2.12 ± 0.17	1.77 ± 0.025
8	3FGL J2356.0+4037	GB6 B2353+4020	359.17	40.66	0.331	27.69	1.55 ± 0.91	91.68	2.04 ± 0.23	1.72 ± 0.040
9	3FGL J1714.1-2029	1RXS J171405.2-202747	258.48	-20.41	-	27.34	2.01 ± 1.11	43.06	1.13 ± 0.88	1.59 ± 0.23
10	Unkn. J2132.43-3416	-	323.21	-34.24	-	25.63	2.45 ± 1.84	3.83	<0.42	-
11	3FGL J2209.8-0450	-	332.44	-4.86	-	25.59	2.60 ± 1.40	27.39	1.37 ± 0.032	1.80± 0.0078
1	3FGL J0730.5-6606	PMN J0730-6602	112.80	-66.00	0.106	19.17	<2.10	102.93	2.59 ± 0.96	1.71 ± 0.13
2	3FGL J1309.3+4304	B3 1307+433	197.21	42.83	0.690	20.10	<1.14	1123.02	15.25 ± 0.23	1.92 ± 0.0067
3	3FGL J1659.0-0142	-	255.23	-1.44	-	15.39	<0.86	86.66	10.44 ± 3.79	2.16 ± 0.13
4	2FGL J1721.5-0718c	-	260.18	-7.20	-	12.95	<4.81	5.17	<13.2	-
5	3FGL J1838.8+4802	GB6 J1838+4802	279.68	48.01	0.300	13.57	<0.91	828.92	10.23 ± 1.06	1.79 ± 0.041
6	Unkn. J0255.43+3334	-	43.90	33.57	-	16.53	<1.20	∼0	<0.068	-
7	Unkn. J0808.43+1645	-	122.19	16.75	-	18.34	<5.08	0.03	<2.16	-
8	Unkn. J1359.3-4019	-	209.86	-40.32	-	22.44	<1.56	0.68	<21.3	-
8	Unkn. J1526.16-0515	-	231.57	-5.26	-	12.36	<1.01	∼0	<0.070	-
10	Unkn. J1626.7-0617	-	246.73	-6.29	-	23.45	<1.69	∼0	<0.011	-
11	Unkn. J1655.52+0052	-	253.99	-0.88	-	14.19	<1.22	35.37	18.90 ± 0.042	2.76 ± 0.00039
12	Unkn. J1902.14+4557	-	285.41	46.06	-	15.53	<2.24	0.67	<0.47	-
13	Unkn. J1903.33+3649	-	285.90	36.82	-	10.47	<1.26	7.13	<63.0	-
14	Unkn. J1907.07-2930	-	286.69	-29.36	-	10.34	<0.49	1.82	<10.9	-
15	Unkn. J1938.09-0350	-	294.55	-3.84	-	12.35	<1.51	1.20	<25.9	-
16	Unkn. J2001.5+0330	-	300.47	3.68	-	16.44	<1.04	∼0	<0.065	-
17	Unkn. J2212.19+8221	-	333.08	82.36	-	20.43	<1.12	∼0	<0.056	-

Table 6.3: Results for sources detected at $E_\gamma \geqslant 100$ GeV with DBSCAN. 'Unkn.' refers to sources that are not present in the 3FGL and z is the redshift of known counterparts. The Test Statistic (TS), flux and Γ were found with follow-up analysis using the published *Fermi* tools. The first 11 sources are those that were found to be significant (TS >25) with the the follow-up analysis. For sources with TS <25, upper limits were calculated for the flux. A binned likelihood analysis has also been applied to the Energy Range 0.1 $>E_\gamma >$100 GeV in order to obtain a power-law spectral index, see Section 6.4.2.

† The two sources RBS 0679 and RBS 0970 were discovered as VHE sources in [44] and [43] respectively but are not in the TeVCat catalogue.

and the TeVCat VHE catalogue (Table 6.2), 21 are also detected using DBSCAN. Of the remaining 28, 11 were found significant with follow up *Fermi* analysis (Table 6.3); 10 of these are in the 3FGL catalogue, which reports fluxes up to only 100 GeV. The positions of these sources can be seen in Figure 6.8.

6.4.1 DBSCAN Performance

To estimate the performance of the DBSCAN algorithm in the case of VHE detections, the concept of *purity* was defined as the number of sources with TS>25 (including the sources already in the TeVCat catalogue) against the total found by the DBSCAN clustering code. The left panel of Figure 6.9 shows the number of sources with TS>25 and TS<25, along with the resultant purity, for all clusters found with DBSCAN using the range of investigated *EPS* values between the 68% and 95% PSF. As can be seen, the number of TS<25 sources found by DBSCAN rapidly increases for *EPS*>0.3°, while there is a marginal increase in the number of sources having TS>25 above the same threshold. It should be noted however that the maximum number of significant sources found by DBSCAN occurs for *EPS*>0.4°. As such, in order to maximise the number of sources with TS>25, with the maximum purity, an *EPS* of 0.4° should be used by DBSCAN. For the remainder of this Chapter, results are based on the DBSCAN results with *EPS*= 0.4°.

To investigate the performance of the likelihood ratio test (LRT) significance, s, in equation 6.2.1, the LRT values for the clusters were compared to the TS values obtained with the *Fermi* Likelihood analysis. In the right panel of Figure 6.9, the LRT vs TS parameter space shows a clear correlation, with a significant quantisation of the LRT distribution for low s values. This quantisation is primarily due to the lack of background events detected with the LAT detector in the $E_\gamma > 100$ GeV energy regime. While this suggests that the use of the LRT to define a DBSCAN cluster as significant results in a large number of false-positive detections, it is noted that the use of a LRT selection criterion of $s > 2.0$ is a conservative cut so as to guarantee the selection of all VHE sources in the sample. As such, while the use of $s > 2.0$ is sub-optimal for selecting VHE candidates with a high purity, Figure 6.9 shows that this allows all VHE sources present within the data set to be found

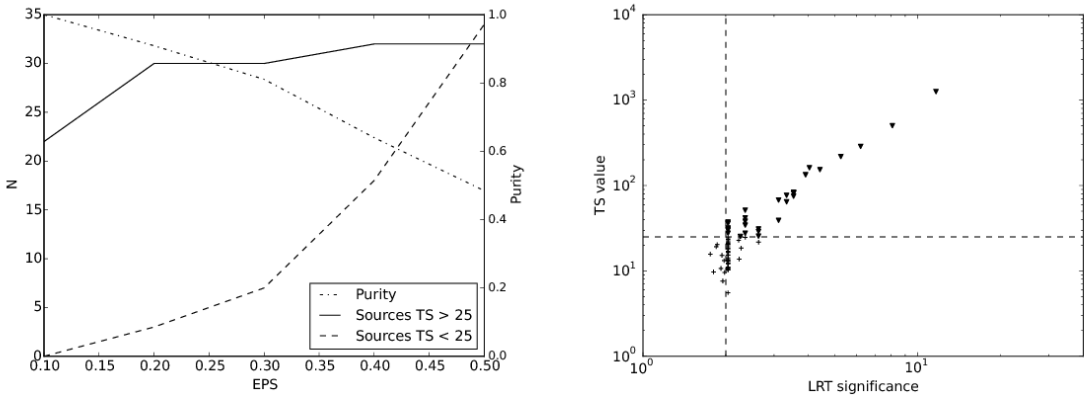


Figure 6.9: (a) Here the effect of different values of *EPS* on the number of sources with TS <25 and TS >25 can be seen. Once over 0.2° the number of significant sources does not greatly increase until 0.4° when one further source is added. However, the number of 'sources' that are not significant continually increases. (b) Comparing the value of LRT and TS for each cluster, the "quantisation" of the LRT due to the small number of signal and background events is clear. The solid triangles indicate the clusters with TS> 25, while the crosses indicate the clusters with TS< 25. The vertical dashed line indicates the LRT cut value, while the horizontal dashed line indicates the TS> 25 cut value.

and thus maximises the number of new sources discovered. Nonetheless, further work should be performed in order to investigate viable alternatives to the LRT that simultaneously maximise both the VHE-detection efficiency **and** the sample purity.

A full understanding of the efficiency of DBSCAN in this application is somewhat more complex, requiring detailed simulations and modelling of the *Fermi* VHE sky. However, estimations of DBSCAN efficiency can be found in [116] where, by simulating a range of false sky maps, they find it possible to achieve efficiencies of up to 96%. This must be treated as an optimistic scenario as it is based on an optimal scan of the *EPS-MinPts* parameter space. It is expected that the efficiency will be much lower in reality due to the assumption of minimal background variation.

Although it is noted that there are still improvements to be made with the DBSCAN method, it is worth noting its ability to perform a quick, unbiased scan for potential "seed" sources in the VHE *Fermi*-LAT sky, which in this study has led to the detection of 9 new VHE sources.

6.4.2 Detected VHE Sources

To investigate the global properties of the *Fermi*-LAT VHE sources detected by the DBSCAN algorithm, a binned likelihood analysis[6] was performed over the energy range 100 MeV to 100 GeV in order to to obtain a reliable model file and fit for eqn. 6.3.2 with higher statistics. The data reduction method for this was the same as described in Section 6.3, but this time using a ROI of 12° centred on the published location of the source, keeping all modelled source parameters within this ROI free and freezing sources within an annulus 12° to 22° around the source of interest. The larger ROI is required to account for the larger PSF found at lower energies. For the analysis, the data were separated into 30 equally-spaced logarithmic energy bins. The resulting fluxes, spectral indices and TS values of the likelihood fits for these objects can be found in Table 6.3. For sources with TS <25, upper limits were calculated from the final fit and no spectral index is quoted.

Out of the 11 sources detected, it is noted that 9 of them are blazars and all, except for 3FGL J1714.1-2029 which is of unknown AGN type, belong to the BL Lac class. The remaining 2 do not have any assigned counterparts. For each source the temporal coincidence of the VHE events was searched for but no evidence was found to suggest that the VHE photons originated in a single event.

The source of unknown type, 3FGL J2209.8-0450, which is a new addition since the 2FGL, is only 54.55" away from the radio source NVSS J220941-045111 (which is also connected to the X-ray object 1RXS J220942.1-045120). The second unassociated source has no known counterpart in the 3FGL (the closest known 3FGL source is the pulsar PSR J2124-3358, at 1.69° from the source) and no clear radio association, although its position is coincident with the galaxy group ESO 403-6. While this source was detected in the 100 GeV to 300 GeV range with a flux of (2.45 $\pm$ 1.84)$\times 10^{-11}$ ph cm^{-2}s^{-1} it appears to have no significant emission in the energy range of 100 MeV to 100 GeV, making this an interesting dark VHE source. More

[6]The tools required for this are almost the same as for the unbinned case shown in Figure 6.6 but require a 3D counts map to be created with GTBIN, the tool GTEXPCUBE2 which in the binned version of GTEXPCUBE and GTSRCMAPS which creates a model counts map for each modelled source (to save time in the likelihood fit)

work needs to be carried out in order to correctly identify counterparts for these sources.

In order to determine the likelihood that any of the unassociated sources with TS <25 are unresolved AGN, a search for any coincidence with BZCAT sources [88] was carried out. No evidence was found of any association within the 95% PSF, suggesting that a large proportion of these clusters arise from fluctuations in the background or from a larger unresolved structure.

As a first check of these results, the spectral index found for each of sources with TS >25 was compared to those published in the 3FGL to look for any change over the last 4 years. There is no evidence of spectral hardening/softening, with the values agreeing within errors.

The spectral index distribution of the sources found using DBSCAN was then compared with the total 3FGL BL Lac population and those which also have ground-based VHE detections. The result of this comparison is shown in Figure 6.10. In order to test whether the different distributions have the same mean and variance, a standard independent 2-sample t test was performed on the DBSCAN sample and each of the spectral index distributions. Having initially set a significance level of 5%, it is apparent that the *Fermi* VHE sources detected with DBSCAN are better represented by the TeVCat BL Lacs, with a P value of 0.368, than the total 3FGL BL Lacs for which a P value of 0.000547 was obtained. It is suggested that the sources which have been detected with VHE emission using DBSCAN analysis, should be within reach of current and future ground based IACTs, provided there are no spectral cut-offs, and should undergo follow up observations.

6.5 Conclusion

In this Chapter, initial results have been presented for the application of the clustering algorithm DBSCAN to 6.25 years of *Fermi*-LAT extragalactic data above 100 GeV. The analsysis returned 49 clusters which were found to be significant using a likelihood ratio test. Of the 28 which are not already known as VHE emitters, published in the TeVCat ground-based catalogue, 11 were found that were significant (TS>25)

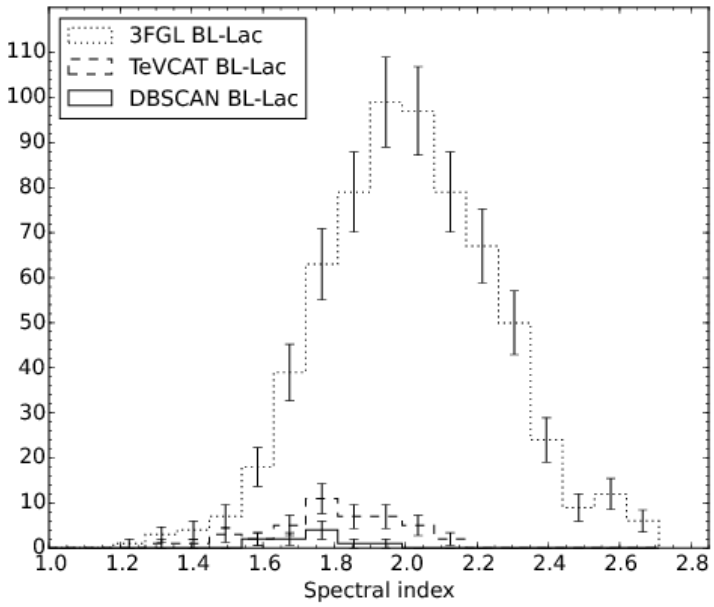


Figure 6.10: Histogram showing the spectral index distribution of the 3FGL and TeVCat BL Lac populations compared to those found in this work. Performing a standard independent 2-sample t test infers that the 11 significant sources in Table 6.3 come from the same distribution as the VHE TeVCat sources.

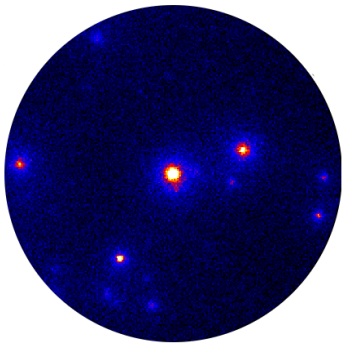
with follow up *Fermi* likelihood analysis. With the 2 sources RBS 0679 and RBS 0970 having previously been detected at E $\geqslant$ 100 GeV [43], [44], this work therefore presents 9 new VHE objects consisting of 7 AGN and 2 unassociated sources.

A consideration has been given to some of the global properties of the new *Fermi* VHE sources. Concerning the spectral indices derived from a fit between 100 MeV and 100 GeV, it is seen that these sources are more similar to the the TeVCat BL Lac sources than to the overall 3FGL BL Lac population. This is therefore taken as a strong indication that these should be observable by current and future ground based IACTs.

As this initial investigation into the performance of DBSCAN returned promising results, it is therefore assumed that improvements to the algorithm and analysis will lead to a robust method for source detection. In the following Chapter, several improvements have been considered in addition to using the improved Pass 8 data set released by the *Fermi* collaboration shortly after the work in this Chapter was completed.

Chapter 7

Source Detection in Sparse Data Sets: Pass 8



In this final Chapter, following the release of updated *Fermi* instrument response functions and reprocessed data, which has been extended up to 3 TeV, the method presented in the previous Chapter has been re-applied with the addition of several improvements. Along with this, the *Fermi*-LAT collaboration released the second catalogue of hard sources and therefore, due to the similarity in goals, a comparison of the DBSCAN performance in the same regime was possible and is presented here.

7.1 Introduction

In summer of 2015 the *Fermi*-LAT collaboration released the long anticipated Pass 8 event characterisation for the LAT data and the associated instrument response functions [29]. While the original motivation for the Pass 8 characterisation of the LAT data was to address the issue of instrumental pile up, often referred to as ghost events[1], detailed Monte Carlo simulations of the LAT detector, combined with improved event reconstruction and background rejection has also led to enhancements in a number of other important detector performance areas. These enhancements include better energy and angular resolution, an increased effective area, an extended observable energy range (around 10 MeV $< E_\gamma <$3 TeV) and better background characterisation resulting in an improved point source sensitivity. All these improvements have led to a larger photon data set in which to search for significant clusters of VHE photons.

In the previous Chapter a clustering analysis of *Fermi*-LAT data in the energy range 100 - 300 GeV and with $|b| >10^\circ$ using the algorithm DBSCAN [65] was presented. Applying this to Pass 7 reprocessed data taken between 4th August 2008 to 28th November 2014, nine new VHE AGN were found. The justification for considering only $>$ 100 GeV data was threefold: reducing background and its associated variations, reducing the size of the data set to improve efficiency while performing a full unbiased and model-independent analysis of the whole sky, and finally increasing the number of known VHE objects that would be promising targets for ground based instruments such as CTA.

This Chapter aims to revisit the clustering work, re-applying DBSCAN to the improved Pass 8 data set including photon energies up to 3 TeV. Additionally, the *Fermi* collaboration has released a catalogue of high energy sources, the second *Fermi*-LAT catalogue of high energy sources (2FHL [11]), from analysing the Pass

[1]In the original reconstruction software for *Fermi*-LAT the presence of signal from non-triggering events which would appear in the reconstruction of a triggered event was not considered. These unexpected signals, referred to as ghost events, act to reduce the accuracy with which an event can be reconstructed. The new Pass 8 software now includes an algorithm to identify and exclude such signals.

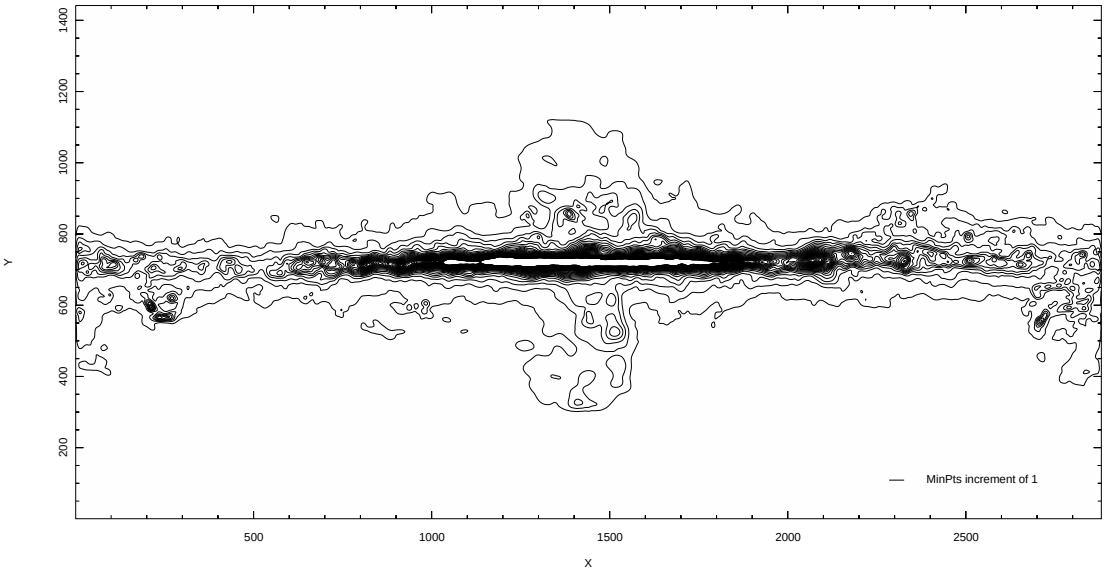


Figure 7.1: Representation of the look-up table for the variable background based on the *Fermi*-LAT Galactic diffuse model. Here the contours represent steps in *MinPts* with a minimum being found at high latitudes. The X and Y scales are in pixel number within the image.

8 data between the energy range 50 GeV to 2 TeV. Therefore it is only logical, considering the similar goals in this clustering work and the 2FHL, to perform a comparative study.

The work presented here is organised in the following way. In Section 7.2 the improvements that have been implemented over the previous work will be explained and in Section 7.3 the method employed to extrapolate the isotropic and Galactic diffuse background models, which do not currently extend up to the full 3 TeV, will be discussed. In Section 7.4 the data preparation and event selection will be specified and the maximum likelihood method for source verification will be summarised in Section 7.5, with the results discussed in Section 7.6. In Section 7.7 a comparison with the 2FHL will be given and a final discussion and conclusion will be presented in Section 7.8.

7.2 Improvements to DBSCAN

The density-based clustering code DBSCAN was described in detail in the previous chapter. In summary, the algorithm considers the number of points contained within a radius *EPS* and starts to build up a cluster if this number exceeds a threshold number of point, defined as *MinPts*, that would constitute an over density above the expected background.

The previous analysis presented made use of the central DBSCAN Python library. However, in order to apply certain improvements in this Chapter a custom implementation was written in C. One of the initial improvements as a consequence of the change was an order of magnitude boost in speed; while not currently a limiting factor at these energies, it provides the possibility to investigate larger data sets in the future. The main improvements where related to tuning the input parameters of DBSCAN, *EPS* and *MinPts* based on a priori information about *Fermi*-LAT, and the development of an improved significance estimator for the clusters found.

7.2.1 Variable DBSCAN Input

The input parameters for DBSCAN can be related to real properties of *Fermi*-LAT. The *EPS*, the radius within which the search for neighbouring points with DBSCAN is carried out, is an analogue for the LAT Point Spread Function (PSF) which varies as a function of energy and conversion type. *MinPts*, or the minimum number of points found within the *EPS* neighborhood, can be related to an over density above the expected background.

The Point Spread Function determines the how well the *Fermi*-LAT instrument can resolve point sources, creating a spread in the localisation. From this, considering the description of the *EPS* parameter, it is clear that one can be related to the other. As previously discussed this was in fact shown in [116], where a detailed analysis of DBSCAN's performance on simulated *Fermi*-LAT data was performed. However, it is noted in [11], that the PSF (as well as the effective area) is only weakly dependent on energy in the E >50 GeV domain (see Figure 7.2). Therefore, the *EPS* has been fixed to the value derived in the previous Chapter, EPS=0.4°.

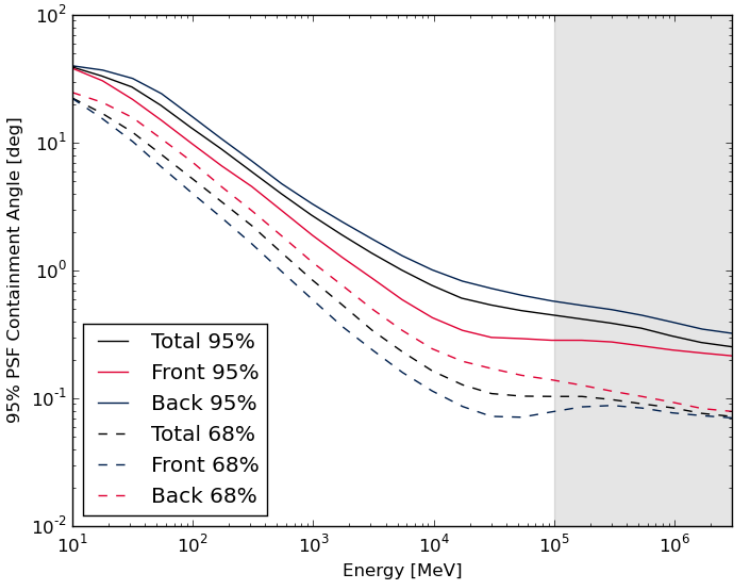


Figure 7.2: The *Fermi*-LAT Pass 8 point spread function for FRONT, BACK and total converting SOURCE class events for 95% and 68% containment.

In future work, where it may become desirable to perform cluster analysis in higher dimensions (e.g. in time or energy) it is foreseen that this may become a more important factor.

Within the DBSCAN *EPS* parameter neighborhood, a minimum number of points is required in order to find and grow a cluster. One of DBSCAN's major perceived downsides is that the algorithm performs poorly when working on data with variable noise, where certain clusters will either be lost or absorbed depending on the initial choice of parameters. However, in this particular application, there exists a relatively good understanding of variations in the background and it is here the *Fermi*-LAT Galactic diffuse background model[2] can be included.

The Galactic diffuse model was created by the *Fermi* collaboration by fitting the observed LAT background to a combination of data from spectral line surveys mapping out HI and CO (as a tracer of H2) and models based on inverse Compton emission, bremsstrahlung and pion decay created by the galaxy modelling software GALPROP [45]. Additional components within the model come from the *Fermi*

[2]Current model used in the work: gll_iem_v06.fits

Bubbles which are believed to be the inverse Compton scattering component of the WMAP Haze [58] and high energy emission from the Loop I excess [48].

In order to create a lookup table based on Galactic longitude and latitude (l,b) that could be used to estimate the *MinPts* parameter, each pixel in the model was integrated between 100 GeV and 3 TeV and then scaled,

$$MinPts(l,b) = \alpha \int_{E_{min}}^{E_{max}} \frac{dF(l,b)_{gal}}{dE} dE, \tag{7.2.1}$$

where α is chosen in such a way that equates the pixel with the least background to the desired initial minimum *MinPts* value. Therefore the *MinPts* parameter will scale from this predefined minimum to a larger value as clusters closer to the Galactic plane are evaluated. In Figure 7.1 the increase in *MnPts* can be seen, where each contour line repents an integer increase in the DBSCAN parameter. Due to the dense nature of the Galactic plane, where we find a large amount of source confusion due to the large number of sources and the high level of background, the clustering analysis is performed on extragalactic data with $|b|$ >10°. Therefore, in this application the adaptive background is likely to only affect clusters within and around the *Fermi* Bubbles.

7.2.2 Poisson significance estimation

In order to estimate the significance of each cluster found by DBSCAN the Likelihood Ratio Test (LRT) [85], as is often used in gamma-ray astronomy, was considered in the last Chapter. However, the number of events found in the majority of identified clusters and in the corresponding background were too small to be used with the LRT (less than 10 events), and instead the use of a Poisson probability estimator was investigated.

To evaluate this, the probability of finding the number of events within the containment radius of the cluster r_c was considered, where r_c is the minimum radius that encompasses all events that belong to each cluster derived from the maximum distance between the furthest points. The number of background events between $2r_c$ and $5r_c$ was then found, or, in cases where very few events were found, the outer radius was increased in iterative steps of 0.001° until at least 5 events were

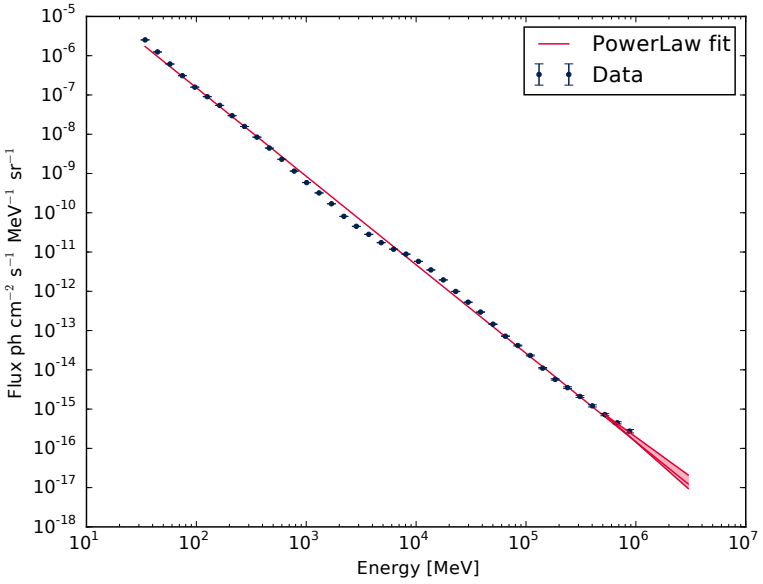


Figure 7.3: Power law fit [red line] to the *Fermi* isotropic background [blue points]. Also shown is an error region in the extrapolated energy range above 878 GeV representing a ±10% change in the spectral index.

contained. The number of background events was then scaled to match the area represented by the containment radius and the probability then calculated with

$$P_{src}(k;\lambda) = 1 - \frac{\lambda^k e^{-\lambda}}{k!}, \tag{7.2.2}$$

where k is the number of source events and λ the expected number of background events. This is simply the probability mass function for a Poisson distribution.

A benefit of using a probability estimator is that it allows a more relaxed initial choice of the clustering parameters. For example, a lower *MinPts* can be adopted which will result in a larger number of clusters, but many of these can later be discarded due to a low probability. For the work presented here, a 0.001% confidence level cut, roughly equating to 4.4σ, has been applied.

7.3 Background Extrapolation

The background models that are currently available from the *Fermi*-LAT team only extend to around 500 GeV while the Pass 8 data is available up to 3 TeV. Hence in

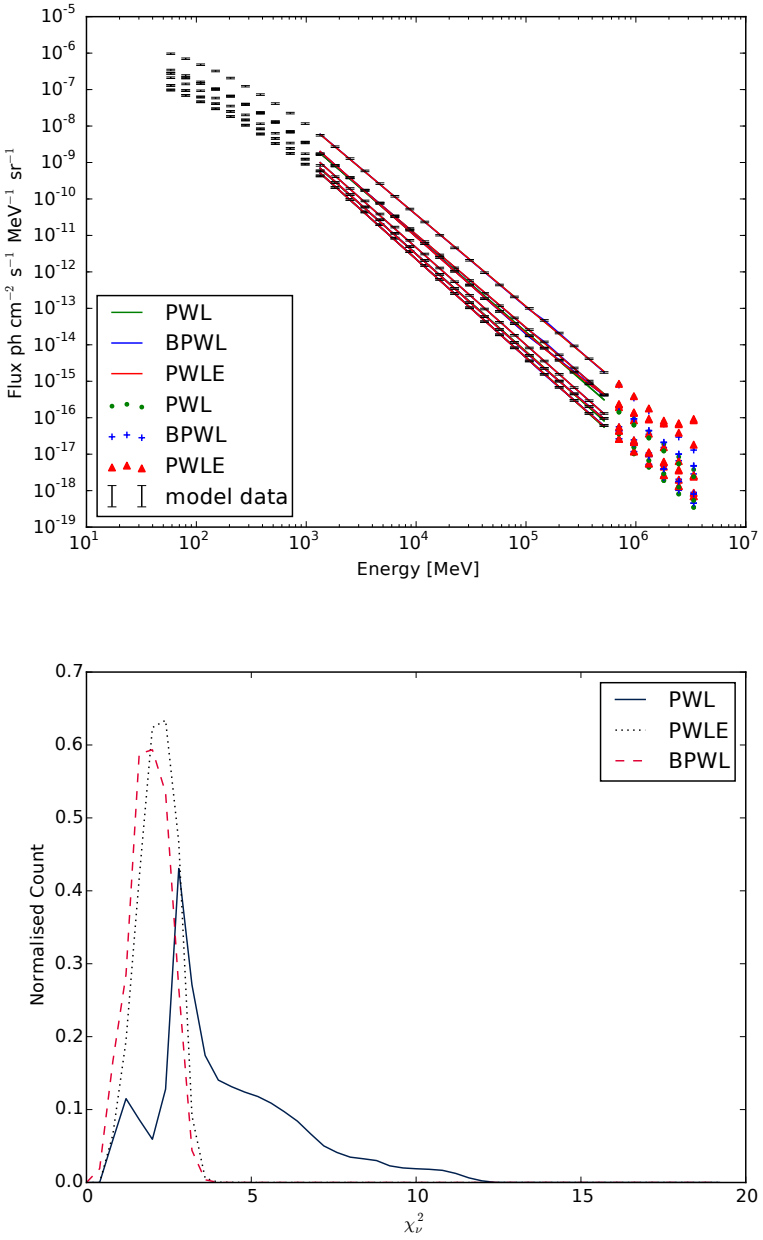


Figure 7.4: Top: A selection of the 2880 × 1441 pixels in the gll_iem_v06.fits Galactic background model and the resulting power law (PWL), broken power law (BPWL) and power law with an exponential cut off (PLWE) fit. Bottom: The distribution of the χ^2_ν fit for each pixel in the Galactic background model.

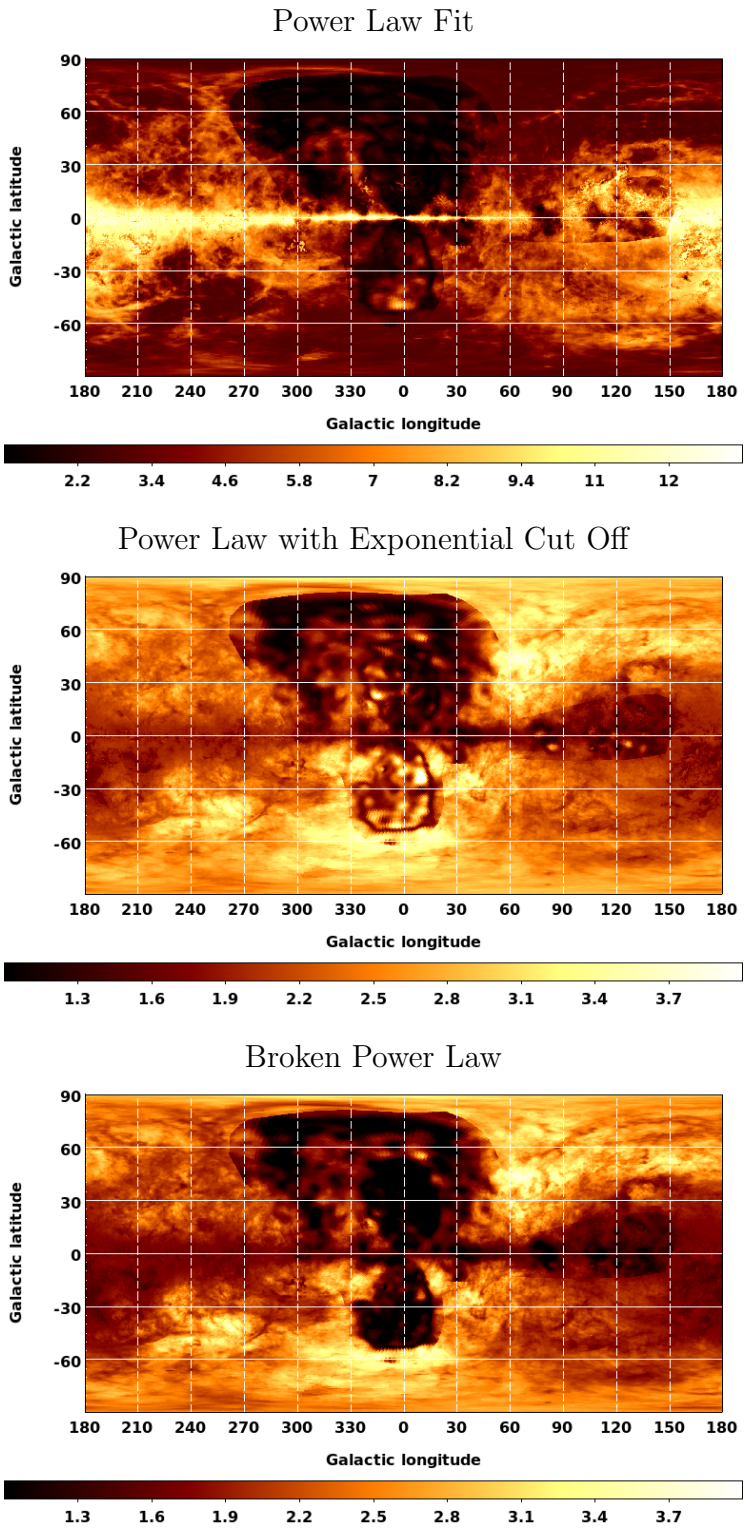


Figure 7.5: Spatial distribution of the reduced χ^2 for different fit functions applied to the Galactic background model. While the pure power law performs poorly, especially for low latitude contributions from the Galactic plane, both the broken power law and the power law with exponential cut off represent the Galactic diffuse well between 1 GeV and 500 GeV

order to analyse this larger data set the Galactic diffuse and isotropic background models need to be extrapolated. However, working with data up to these energies in not advised by the *Fermi* collaboration, therefore results for data up to 500 GeV as well as up to 3 TeV will be presented.

The isotropic background model[3], which was created from a fit to the all sky extragalactic emission not covered by the Galactic diffuse model, consists only of a single energy spectrum extending up to 877.9 GeV. In order to obtain an extrapolation, the data were assumed to follow a power law which can be expressed in the form

$$\frac{dN}{dE}_{iso} = (4.315) \times 10^{-3} \times E^{-2.239}. \tag{7.3.3}$$

This is then used to extrapolate points up to 3 TeV where the flux falls to a level of $1.36{\times}10^{-17}$ ph cm^{-2} s^{-1}. The contribution from the extrapolated fit, between 877.9 GeV and 3 TeV, amounts to 5.4% of the total flux above 100 GeV. Introducing a 10% error in the spectral index for estimating the additional contribution results in a 0.87% to 1.65% uncertainty (See Figure 7.3) which is assumed to be negligible.

The most recent Galactic diffuse model, gll_iem_v06.fits, is a 2880 × 1441 pixel image, with each pixel consisting of thirty logarithmic energy bins, ranging from 58 MeV to 513 GeV. In order to extrapolate this model to 3 TeV, the method used for the isotropic fit was applied, but for each pixel in the image above 1 GeV where the spectrum flattens (see e.g. left panel in Figure 7.4). To determine the model that best represents the Galactic diffuse emission, the data were fitted with a power law (PWL), a broken power law (BPWL) and a power law with an exponential cut off (PWLE). For each fit the reduced χ^2 was calculated and the total distribution can be seen in the upper panel of Figure 7.4. From this it is clear that a pure power law is not the best fit and that the data set is better represented by a BPWL or PWLE. As the Galactic model consists of a combination of models, it is interesting to see the spatial distribution of the fit quality. Therefore in Figure 7.5 the reduced χ^2 has been shown as a function of its spatial position. It can be seen that all models

[3]Current file iso_P8R2_SOURCE_V6_v06.txt available from *Fermi* website.

fit the *Fermi* bubbles and WMAP Haze relatively well. The Galactic emission is poorly fitted by the PWL but is well represented by the PWLE and BPWL.

7.4 Data

The pre-processing of the data was similar to that presented in the previous Chapter. For the Pass 8 data set, both FRONT and BACK converting *Fermi*-LAT events were selected for the first 85 months of data (Mission Elapsed Time: 239557417 to 459745155). In order to process the data, the *Fermi* tool GTSELECT is used to select only FRONT and BACK converting SOURCE (evclass 128) type events above 100 GeV and a zenith angle cut[4] applied, here taken to be 100°.

The second step is the selection of good time intervals with the GTMKTIME tool, for which the recommended filter expression "(DATA_QUAL>0) && (LAT_CONFIG==1)" is provided and removes further removes sub-optimal data[5]. Lastly data with $|b|$ <10° were removed from the file and the clustering algorithm was applied, returning a list of potential sources.

7.5 Maximum Likelihood Follow-up

As the probability estimate for each cluster does not take into account any information about the *Fermi*-LAT instrument performance, each potential source is further investigated using the maximum likelihood analysis which is part of the standard *Fermi* tools. The method used is the same as presented in Chapter 6 Section 6.3. In summary, data within 5° of each cluster were selected. Exposure and livetime were calculated for each ROI. Each point source was modelled with a power law and the total model included all sources within 15° along with the extended background models described in Section 7.3. From this a maximum likelihood analysis was carried out to obtain a best fit model for each cluster. Finally we derive the integrated

[4]This cut is looser than the recommended 90° but more stringent than the 105° used for the event selection in the 2FHL (discussed further in section 7.7).

[5]Note that, for Pass 8 analysis, it is no longer advised to use ROCK_ANGLE <52°.

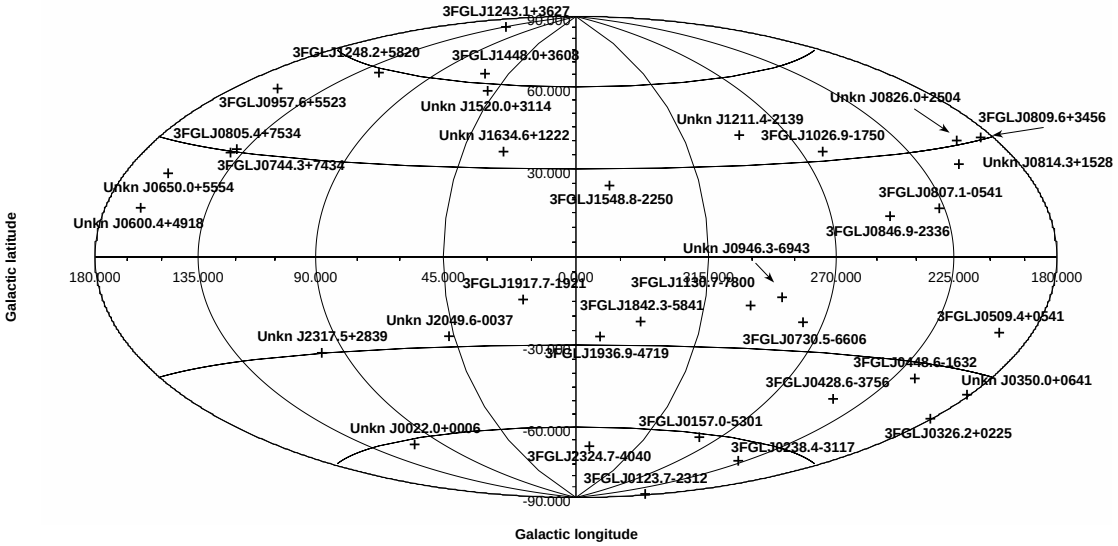


Figure 7.6: Clusters found with a TS >25 in the 100 GeV to 3 TeV range.

flux and TS after applying the GTFINDSRC tool which is used to determine the best fit position of the clusters Right Assention and Declination. The resulting significant clusters can be found in Table 7.1.

7.6 100 GeV - 3 TeV Cluster Results

Using 85 months of *Fermi*-LAT VHE data (above 100 GeV) to which the clustering algorithm DBSCAN is applied, with an input *EPS* of 0.4° and a *MinPts* scaling up from 2 depending on the spatial model of the Galactic diffuse model, a total of 232 clusters were found to pass a probability cut. These consisted of 120 that are already known in the 3FGL, 107 in the 2FHL and 43 in the TeVCat catalogue (lists are not exclusive). The remaining 106 clusters with no obvious association are classed as unknown (Unkn.). After applying a full unbinned maximum likelihood analysis, 36 clusters were found with a TS >25 of which 12 are not associated with any 3FGL sources. Comparing to the 2FHL it is seen that 24 of the sources correspond to published sources, 22 of which are 3FGL objects and 2 Unkn. These results can be seen in Table 7.1 and Figure 7.6 including results using both the PWLE and BPWL

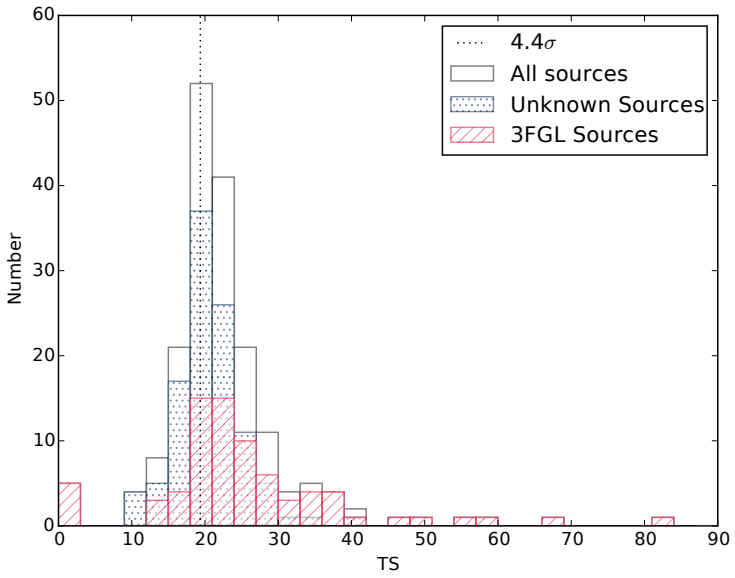


Figure 7.7: Distribution of the TS values for all clusters, showing that the majority of clusters are found with a TS >16, roughly 4σ. Sources that are present in the 3FGL have a higher TS value on average, which is not surprising as their increased flux will have led to their discovery.

Galactic diffuse background models. For any sources that fell below a TS of 25 in one or more of the energy ranges, while upper limits are usually calculated, this becomes increasingly difficult due to the very low statistics and therefore no results are shown.

7.6.1 Revisiting the Probability Cut

Recall in Section 7.2.2 that a probability cut based on equation 7.2.2 was applied to each cluster found by the DBSCAN algorithm. Using a cut level of 0.001%, roughly equal to 4.4σ, a total of 232 clusters was found. While this may seem like a large number of potential sources, it is worth noting that the distribution of TS values obtained centres around this cut value and that almost 90% of sources have a significance of over 4σ (see Figure 7.7). While there is no intention to claim these as sources, this increases confidence in the DBSCAN method.

			RA [deg]	Dec [deg]	TS			Flux $\times 10^{-11}$ ph cm^{-2} s^{-1}		
	Name	2FHL	J2000	J2000	500 GeV	3 TeV PWLE	3 TeV BPWL	500 GeV	3 TeV PWLE	3 TeV BPWL
1	Unkn J0022.0+0006	Y	5.50	0.10	40.18	37.39	37.39	1.91 ± 0.88	1.84 ± 0.87	1.84 ± 0.87
2	3FGL J0123.7-2312	Y	20.96	-23.18	36.61	39.46	39.47	1.11 ± 0.64	1.11 ± 0.64	1.11 ± 0.64
3	3FGL J0157.0-5301	Y	29.20	-53.03	26.72	29.56	29.37	0.92 ± 0.56	0.93 ± 0.56	0.92 ± 0.56
4	3FGL J0238.4-3117	Y	39.63	-31.28	32.32	29.80	29.59	1.09 ± 0.64	1.12 ± 0.65	1.11 ± 0.65
5	3FGL J0326.2+0225	Y	51.58	2.46	32.90	32.54	32.54	1.05 ± 0.63	1.04 ± 0.49	1.04 ± 0.49
6	Unkn J0350.0+0641	N	57.50	6.69	35.71	32.57	32.22	1.08 ± 0.63	1.04 ± 0.62	1.04 ± 0.62
7	3FGL J0428.6-3756	Y	67.18	-37.93	27.51	28.84	28.84	1.20 ± 0.69	0.63 ± 0.46	0.63 ± 0.46
8	3FGL J0448.6-1632	N	72.22	-16.50	28.75	29.31	29.09	0.76 ± 0.54	0.78 ± 0.55	0.77 ± 0.54
9	3FGL J0509.4+0541	Y	77.46	5.76	27.51	26.13	26.88	1.20 ± 0.69	1.21 ± 0.69	1.19 ± 0.69
10	Unkn J0600.4+4918	N	90.17	49.30	26.90	25.45	25.45	0.65 ± 0.44	0.57 ± 0.41	0.57 ± 0.41
11	Unkn J0650.0+5554	N	102.53	55.91	29.78	27.58	27.58	0.98 ± 0.52	0.94 ± 0.50	0.94 ± 0.50
12	3FGL J0730.5-6606	Y	112.76	-66.00	36.46	36.29	36.29	0.90 ± 0.52	0.89 ± 0.51	0.89 ± 0.51
13	3FGL J0744.3+7434	Y	116.20	74.58	32.94	31.93	31.92	0.89 ± 0.48	0.83 ± 0.47	0.87 ± 0.47
14	3FGL J0805.4+7534	Y	121.28	75.57	34.26	34.17	33.87	0.90 ± 0.46	0.90 ± 0.46	0.91 ± 0.46
15	3FGL J0807.1-0541	N	121.68	-5.93	26.98	26.91	27.04	0.78 ± 0.55	0.78 ± 0.55	0.78 ± 0.55
16	3FGL J0809.6+3456	Y	122.39	34.97	35.62	35.93	35.77	0.94 ± 0.54	0.95 ± 0.55	0.93 ± 0.55
17	Unkn J0814.3+1528	N	123.63	15.49	30.18	31.64	31.40	1.02 ± 0.64	1.49 ± 0.81	1.48 ± 0.81
18	Unkn J0826.0+2504	N	126.52	25.07	13.36	26.42	26.23	...	0.73 ± 0.52	0.74 ± 0.53
19	3FGL J0846.9-2336	Y	131.69	-23.58	52.31	49.76	49.75	1.78 ± 0.80	1.78 ± 0.81	1.79 ± 0.81
20	Unkn J0946.3-6943	N	146.61	-69.72	25.31	25.34	25.15	0.54 ± 0.39	0.53 ± 0.39	0.54 ± 0.39
21	3FGL J0957.6+5523	Y	149.48	55.38	27.24	27.57	27.68	0.64 ± 0.44	0.64 ± 0.44	0.64 ± 0.44
22	3FGL J1026.9-1750	Y	156.79	-17.88	24.97	25.85	25.11	...	0.72 ± 0.51	0.73 ± 0.52
23	3FGL J1130.7-7800	Y	172.54	-78.02	33.95	33.70	34.45	0.97 ± 0.52	0.96 ± 0.51	0.97 ± 0.52
24	Unkn J1211.4-2139	N	182.92	-21.65	27.41	25.66	25.62	0.72 ± 0.51	0.71 ± 0.51	0.70 ± 0.51
25	3FGL J1243.1+3627	Y	190.64	36.32	25.06	25.39	23.71	0.89 ± 0.54	0.89 ± 0.54	...
26	3FGL J1248.2+5820	Y	192.11	58.34	33.89	31.11	32.54	0.91 ± 0.51	0.89 ± 0.50	0.87 ± 0.49
27	3FGL J1448.0+3608	Y	222.04	36.13	26.09	26.07	25.97	0.58 ± 0.42	0.60 ± 0.43	0.60 ± 0.43
28	Unkn J1520.0+3114	N	230.01	31.25	14.31	28.34	28.32	...	0.95 ± 0.63	0.96 ± 0.63
29	3FGL J1548.8-2250	Y	237.21	-22.80	38.28	36.22	35.69	1.67 ± 0.76	1.72 ± 0.79	1.72 ± 0.79
30	Unkn J1634.6+1222	N	248.74	12.38	15.99	27.32	27.32	...	1.17 ± 0.70	1.17 ± 0.70
31	3FGL J1842.3-5841	Y	280.55	-58.70	42.26	45.36	45.52	1.20 ± 0.62	1.21 ± 0.63	1.22 ± 0.63
32	3FGL J1917.7-1921	Y	289.44	-19.36	72.79	82.39	82.40	2.52 ± 0.97	2.90 ± 1.04	2.90 ± 1.04
33	3FGL J1936.9-4719	Y	294.24	-47.36	27.53	27.18	27.16	0.79 ± 0.54	0.76 ± 0.53	0.76 ± 0.53
34	Unkn J2049.6-0037	N	312.48	-0.62	25.65	25.29	25.07	1.14 ± 0.67	1.14 ± 0.68	1.17 ± 0.69
35	Unkn J2317.5+2839	Y	349.47	28.65	26.75	37.39	37.39	0.68 ± 0.48	1.84 ± 0.87	1.84 ± 0.87
36	3FGL J2324.7-4040	Y	351.19	-40.70	29.14	29.00	29.00	0.99 ± 0.59	0.96 ± 0.58	0.98 ± 0.59

Table 7.1: Results for sources detected at E $\geqslant$ 100 GeV with DBSCAN. 'Unkn.' refers to sources that are not present in the 3FGL. The Test Statistic (TS) and flux were found with follow-up analysis using the published *Fermi* tools. Results using a Galactic diffuse model extrapolated with a power law with an exponential cut off (PWLE) and a broken power law (BPWL) are shown along with results obtained with published background models (limited to 500 GeV).

	Name	RA [deg] J2000	Dec [deg] J2000	Association	Type	z	Var	TS	Flux $\times 10^{-9}$ ph cm^{-2}s^{-1}	Γ
1	Unkn J0022.0+0006	5.50	0.10	...	...	...	...	17.6	<1.73	...
2	3FGL J0123.7-2312	20.96	-23.18	1RXS J012338.2-231100	BLL	0.404	51.7	286.6	3.85±0.06	1.81± 0.0106
3	3FGL J0157.0-5301	29.20	-53.03	1RXS J015658.6-530208	BLL	...	62.6	334.9	2.79±0.29	1.68± 0.0422
4	3FGL J0238.4-3117	39.63	-31.28	1RXS J023832.6-311658	BLL	0.233	47.5	714.1	6.02±0.69	1.79± 0.0497
5	3FGL J0326.2+0225	51.58	2.46	1H 0323+022	BLL	0.147	59.1	314.3	6.79±0.13	1.36± 0.0068
6	Unkn J0350.0+0641	57.50	6.69	...	...	...	...	13.0	<1.13	...
7	3FGL J0428.6-3756	67.18	-37.93	PKS 0426-380	BLL	1.111	2951.5	64088.0	23.23±3.85	lgp
8	3FGL J0448.6-1632	72.22	-16.50	RBS 0589	BLL	...	56.1	397.0	5.36±0.83	1.84± 0.0636
9	3FGL J0509.4+0541	77.46	5.76	TXS 0506+056	BLL	...	285.3	5577.0	68.71±0.73	2.05± 0.0054
10	Unkn J0600.4+4918	90.17	49.30	...	...	...	...	29.3	90.82±0.36	2.75± 0.0000
11	Unkn J0650.0+5554	102.53	55.91	...	...	...	...	0.0	<3.50	...
12	3FGL J0730.5-6606	112.76	-66.00	PMN J0730-6602	BCU	0.106	41.1	162.8	3.40±0.60	1.83± 0.0628
13	3FGL J0744.3+7434	116.20	74.58	MS 0737.9+7441	BLL	0.314	67.6	305.8	5.41±0.14	2.03± 0.0101
14	3FGL J0805.4+7534	121.28	75.57	RX J0805.4+7534	BLL	0.121	95.4	1253.0	9.42±0.12	1.88± 0.0053
15	3FGL J0807.1-0541	121.68	-5.93	PKS 0804-05	BLL	...	94.6	900.2	23.13±0.24	2.13± 0.0042
16	3FGL J0809.6+3456	122.39	34.97	B2 0806+35	BLL	0.083	47.1	79.3	2.03±0.12	1.91± 0.0190
17	Unkn J0814.3+1528	123.63	15.49	...	...	...	...	30.1	10.55±0.56	4.46± 0.1294
18	Unkn J0826.0+2504	126.52	25.07	...	...	...	...	0.0	<1.00	...
19	3FGL J0846.9-2336	131.69	-23.58	PMN J0847-2337	BLL	0.059	64.8	548.0	10.56±0.12	1.96± 0.0043
20	Unkn J0946.3-6943	146.61	-69.72	...	...	...	...	8.0	<7.98	...
21	3FGL J0957.6+5523	149.48	55.38	4C +55.17	FSRQ	0.899	42.4	25137.0	83.38±0.10	lgp
22	3FGL J1026.9-1750	156.79	-17.88	1RXS J102658.5-174905	BLL	0.26689	152.3	826.0	13.55±0.13	1.96± 0.0446
23	3FGL J1130.7-7800	172.54	-78.02	...	...	...	46.6	164.0	1.41±0.39	1.51± 0.1006
24	Unkn J1211.4-2139	182.92	-21.65	...	...	...	...	14.0	<12.0	...
25	3FGL J1243.1+3627	190.64	36.32	Ton 116	BLL	1.066	58.2	2312.0	10.13±0.63	1.72± 0.0249
26	3FGL J1248.2+5820	192.11	58.34	PG 1246+586	BLL	0.847	65.2	8709.0	35.60±0.80	1.85± 0.0115
27	3FGL J1448.0+3608	222.04	36.13	RBS 1432	BLL	1.508	92.9	1505.0	9.63±0.79	1.79± 0.0364
28	Unkn J1520.0+3114	230.01	31.25	...	...	...	...	30.9	12.88±0.58	5.66± 1.202
29	3FGL J1548.8-2250	237.21	-22.80	PMN J1548-2251	BLL	...	49.1	515.5	9.03±0.22	1.86± 0.0103
30	Unkn J1634.6+1222	248.74	12.38	...	...	...	...	6.4	<1.47	...
31	3FGL J1842.3-5841	280.55	-58.70	1RXS J184230.6-584202	BCU	...	55.3	86.0	1.62±0.19	1.75± 0.0408
32	3FGL J1917.7-1921	289.44	-19.36	1H 1914-194	BLL	0.137	72.9	2111.0	22.28±0.34	1.85± 0.0062
33	3FGL J1936.9-4719	294.24	-47.36	PMN J1936-4719	BLL	0.265	62.5	347.6	1.33±0.07	1.74±0.0169
34	Unkn J2049.6-0037	312.48	-0.62	...	...	...	...	0.0	<3.70	...
35	Unkn J2317.5+2839	349.47	28.65	...	...	...	...	74.9	4.79±0.26	2.15± 0.0149
36	3FGL J2324.7-4040	351.19	-40.70	1ES 2322-409	BLL	0.174	41.3	791.6	6.33±0.74	1.74± 0.0453

Table 7.2: Maximum likelihood results for all clusters in 100 MeV to 100 GeV energy range. For sources that have a TS <25, upper limits to the flux are presented and no spectral index is shown. In this table, known associations of 3FGL sources are shown along with their source classification (BLL = BL Lac type blazar, BCU= active galaxy of uncertain type and FSRQ = flat spectrum radio quasar type blazar), redshift if known and published variability index (where variability index is described in Section 7.6.2).

7.6.2 E <100 GeV and Variability

In order to obtain a better understanding of the sources found using clustering analysis, a binned likelihood analysis over the energy range 100 MeV to 100 GeV is also carried out as in Chapter 6. To summarise, the data selection is as before (Section 7.4), with the exception of a larger ROI of 12° to account for the increased PSF at lower energies, including sources in the model file out to a further 10° with parameters frozen to those in the 3FGL.

By performing a likelihood fit with greater statistics, improved fit parameters for these sources are determined. The low energy fit results can be seen in Table 7.2. In this table any associations of objects already known in the lower energy band, along with object type and redshift, are included (if known). It is noted, that while most of the sources are BL-Lac types, the Flat Spectrum Radio Quasar (FSRQ), 4C +55.17 is also found along with two active galaxies of uncertain type (BCU) which may prove to be BL-Lacs.

An important factor to take into account when considering these sources for follow up observations with ground based IACTs is whether or not their discovery depends on flaring events or strong variability over the last 85 months. As an indication, for known 3FGL sources, the published variability index is presented, where the variability is defined as

$$\mathrm{Var} = \sum_{\mathrm{month}} -2\ \ln(L_{\mathrm{month}}/L_{\mathrm{yr}}), \tag{7.6.4}$$

where L_{month} is the likelihood for the flux fitted in a month and L_{yr} is the likelihood of a flat lightcurve fitted over the full catalogue interval. The resulting sum over each month interval gives an indication of the source variability for month-long time scales. This does not take into account shorter time scales and therefore, as previously stated, is used only as an indication. Here, a value greater than 72.44 equates to a 99% confidence limit that the source is variable.

In addition to the variability index, evidence for temporal clustering of the VHE events was also considered. By investigating the distribution of the time between event pairs for all sources detected, evidence of clustering should become apparent

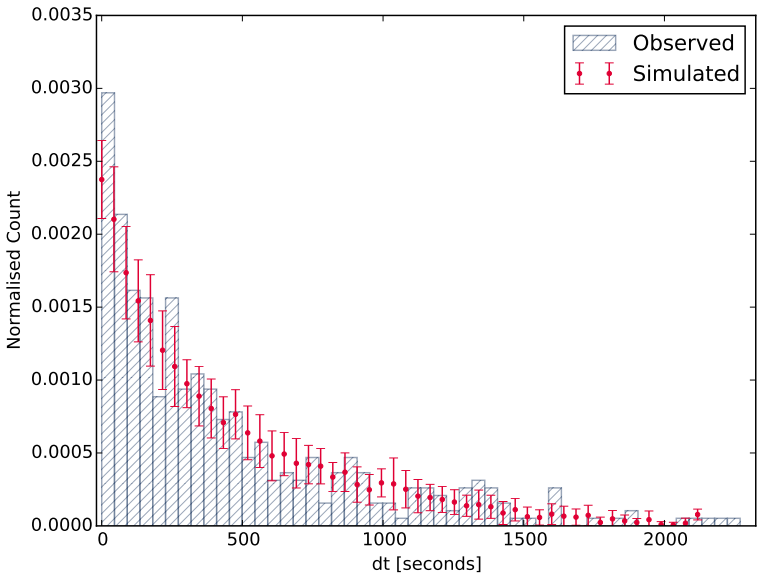


Figure 7.8: Distribution of time between events for all clusters found passing the probability cut (blue hatched histogram) compared to a simulated data set of random arrival times (red errorbars). If the VHE events originated from single flaring events an excess at low dt would be expected.

as an excess at small time separations. In Figure 7.8, the difference in time for consecutive pairs of events within each source data set, for all clusters found passing the probability, cut can be seen as the hatched histogram. The events for each cluster were then randomly assigned an arrival time within the 85 months of observation and the process was repeated multiple times to obtain the simulated data set. From this it appears that there is no strong evidence that the events originated from single bright flares.

7.6.3 Unknown sources and possible associations

Out of the 36 clusters found to be significant in the 100 GeV to 3 TeV range, 24 were spatially coincident with existing 3FGL sources, leaving 12 sources which are classified with the prefix unknown (Unkn). Two of these, Unkn J0022.0+0006 and Unkn J2317.5+2839 have been detected in the 50 GeV to 2 TeV range and are part of the 2FHL, with the former associated with 5BZGJ0022+0006 (BL Lac with prominent galaxy emission) and the latter unassociated. For the remaining

10, possible associations based on source locations from the astronomical database Simbad[6] are considered (the sources and surrounding objects known in Simbad can be seen in Figure 7.9 and 7.10).

Unkn J0022.0+0006

Only significant above 100 GeV, the object is directly coincident with the BL Lac source RX J0022.0+0006 with a redshift of 0.306 at a distance of 44.54 arcseconds from the cluster location. This source is known in the 2FHL catalog as 2FHL J0022.0+0006 and is associated with the BL Lac AGN 5BZGJ0022+0006 with a redshift of 0.306.

Unkn J0350.0+0641

Found with a reasonable significance in all but the lower energy range, this source is spatially coincident with the BL Lac object 2MASS J03495785+0641264 [120] at an angular distance of 31.97 arcseconds. There are also a number of galaxies which form part of a galaxy cluster in the vicinity.

Unkn J0600.4+4918

Found to be significant within each energy range although only slightly over a TS of 25. There appears to be little evidence of possible counterparts close to the source position.

Unkn J0650.0+5554

Significant in all but the low energy range. Apart from a high proper motion star at the centre of the field of view there are two nearby X-ray sources 1RXS J064947.8+555232 and 1RXS J065009.9+555827 at 188.94 and 234.95 arcseconds from the centre respectively. Neither source has any associated classification.

[6] http://simbad.u-strasbg.fr/simbad/

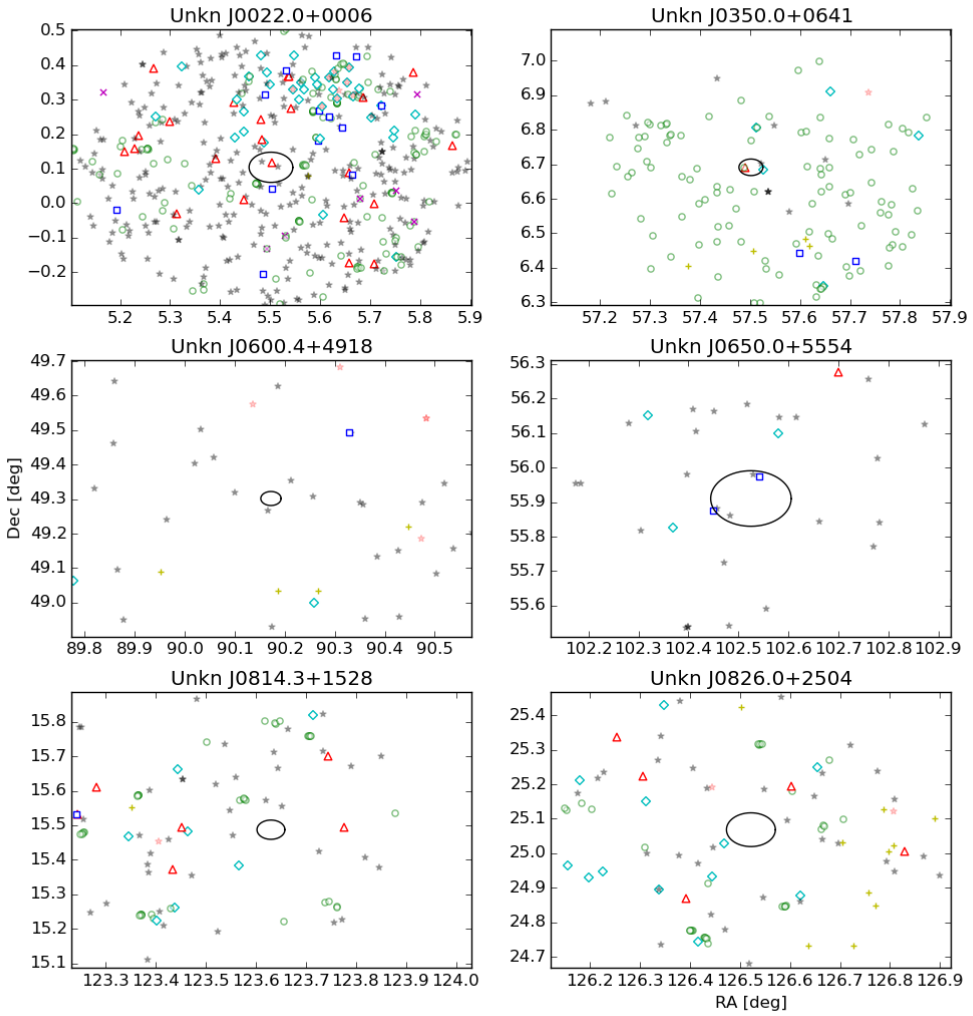


Figure 7.9: Position of Unkn. sources and their 1 σ positional uncertainty (black circle) shown against known sources in the Simbad catalogue, where green circles are galaxies, red triangles are AGN, blue squares are X-ray sources, cyan diamonds are radio objects, yellow crosses are infra red objects, magenta x are Super nova remnants or pulsars and black stars are stars.

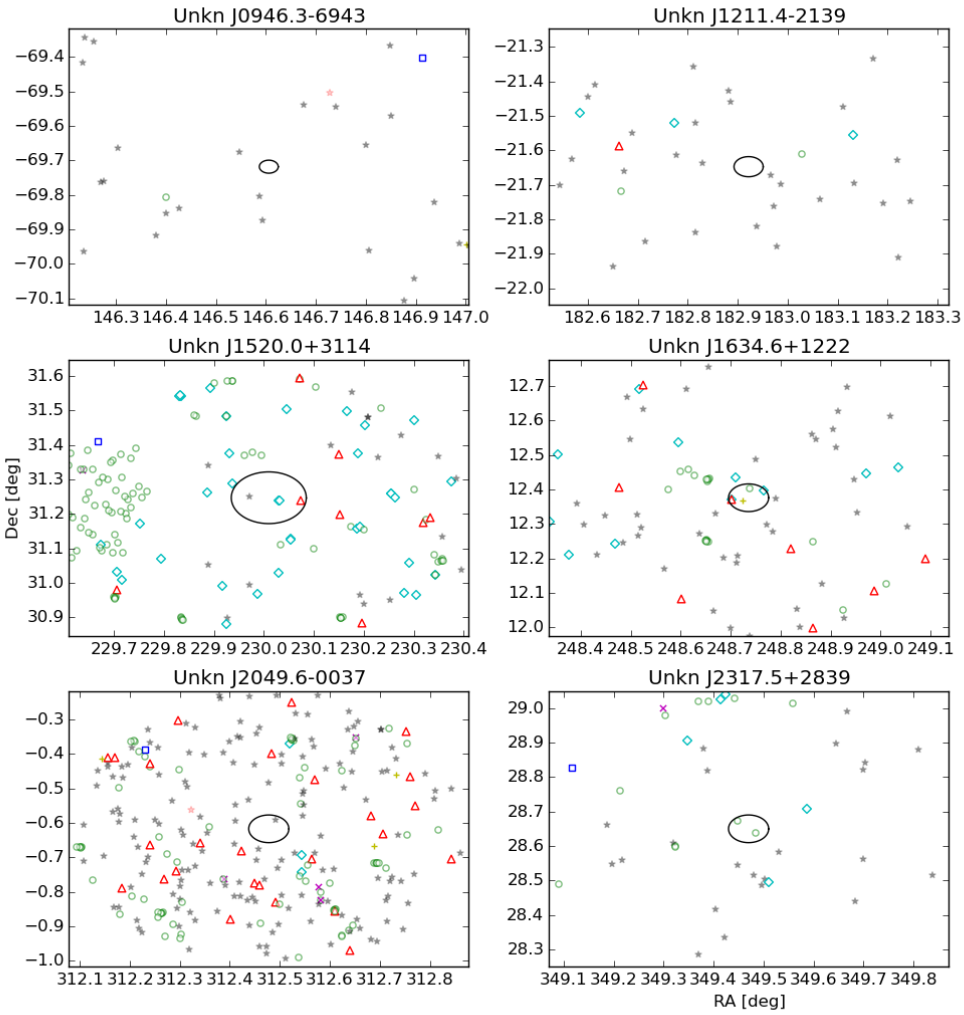


Figure 7.10: See Figure 7.9.

Unkn J0814.3+1528

Significant in all energy ranges, also with a bright star at the centre of the field of view. Nearest possible associations are the X-ray source RX J0254.6+1525.1 of unknown type 163.84 arcseconds away from the centre, the Quasar 2MASS J02543634+1525443 198.22 arcseconds from the centre and a group of galaxies at 488.16 arcseconds from the centre.

Unkn J0826.0+2504

Only found in the highest energy range, 100 GeV - 3 TeV, closest associations are the radio source NVSS J082552+250138 of unknown type at 228.49 arcseconds distance, a cluster of galaxies [SPD2011] 54642 at 458.16 arcseconds and the quasar [VV2006] J082624.6+251142 at 525.85 arcseconds from the centre.

Unkn J0946.3-6943

Significant in all but the low energy range, there appears to be very little in the way of possible associations, with the closest being the X-ray source 1RXS J094420.5-694844 of unknown type 730.96 arcseconds away.

Unkn J1211.4-2139

Again significant in all but the low energy range, with a bright star in centre of the field of view. Nearest possible associations are the radio source NVSS J121231-213315 of unknown type at 758.05 arcseconds and a possible AGN 2MASS J12103884-2135167 904.44 arcseconds away.

Unkn J1520.0+3114

Found to be significant in all but the 100 GeV - 500 GeV range, this source coincides with a multitude of radio sources in the FIRST radio survey [123]. There also exists the quasar [VV2006] J152036.6+311151 at 222.63 arcseconds from the centre and the AGN 2MASX J15203610+3122258 with redshift z=0.10675 at 506.21 arcseconds from the centre of the field of view.

Unkn J1634.6+1222

Only significant in the highest energy range, 100 GeV to 3 TeV. The region of interest includes a large number of galaxies and radio sources due to it being covered by the Sloan Digital Sky Survey [124]. Most notable are the quasars SDSS J163448.25+122215.3 and SDSS J163516.70+121340.6 at a distance of 140.72 and 614.76 arcseconds and with redshifts of z=1.03 and 1.79 respectively.

Unkn J2049.6-0037

Significant in all but the lowest energy range. Closest possible associations is the BL Lac type blazar 1RXS J204921.6-003930 at 519.54 arcseconds and a redshift of 0.25684.

Unkn J2317.5+2839

Significant in all energy bands, but there are no clear associations, with the closest counterparts being the radio sources NVSS J231820+284232 at 423.32 arcseconds and GB6 B2315+2813 at 573.83 arcseconds. This source is known in the 2FHL catalogue as 2FHL J2317.8+2838 but has no known association

7.7 Comparison with 2FHL

Shortly after the release of the Pass 8 reprocessed data, the *Fermi* collaboration released the second *Fermi*-LAT catalogue of hard sources, the 2FHL [11]. Considering the similarity in goals of the 2FHL and this work it is important to compare the relative performance. The 2FHL was created using 80 months (MET 239557417 to 449855019, August 2008 to April 2015)[7] of Pass 8 SOURCE class data between 50 GeV and 2 TeV. For their data reduction they used a zenith cut of 105°, which has also been adopted for this analysis.

[7]This is the time range defined in [11] and on the 2FHL website, however due to the detection of 3FGL sources shown in Section 7.7.1 this has been brought into question (past catalogues provide seed sources for later catalogues). In the header of the 2FHL fits data file, the defined time range equated to only 78 months. This discrepancy is currently under investigation.

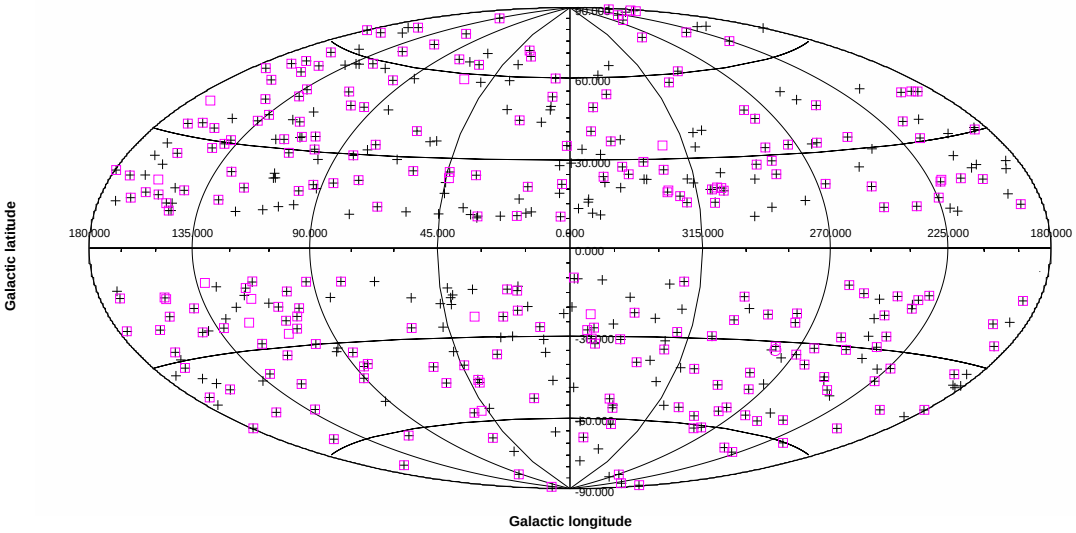


Figure 7.11: Clusters found in the 50 GeV to 2 TeV range (black cross) compared to the 2FHL source catalogue (magenta squares).

For the clustering analysis, the same input parameters as the previous data set were chosen, but requiring a minimum number of points of 3 due to the increased data set size, once again, only events with $|b|>10$ were chosen. For clusters found with a probability passing a 0.001% confidence limit we follow the same procedure as described in Section 7.5. Due to the similarity in results from the different background models presented in the 100 GeV to 3 TeV range, only results using a power law with exponential cut of have been shown.

7.7.1 Cluster results

Running the clustering analysis on the same data set as the 2FHL a total of 422 clusters passing the probability cut are found. Of the 257 possible 2FHL sources with $|b|>10$, 237 matching clusters are found to be spatially coincident, resulting in a 92.2% success rate. Out of the 131 with no 2FHL counterpart, 15 were found to be significant following a full likelihood analysis; these are listed in Table 7.3. Unlike the 3 TeV results, the number of sources that are not in the 2FHL above 4σ make up only 51.9% of the sources (see left panel of Figure 7.12). It is worth noting that

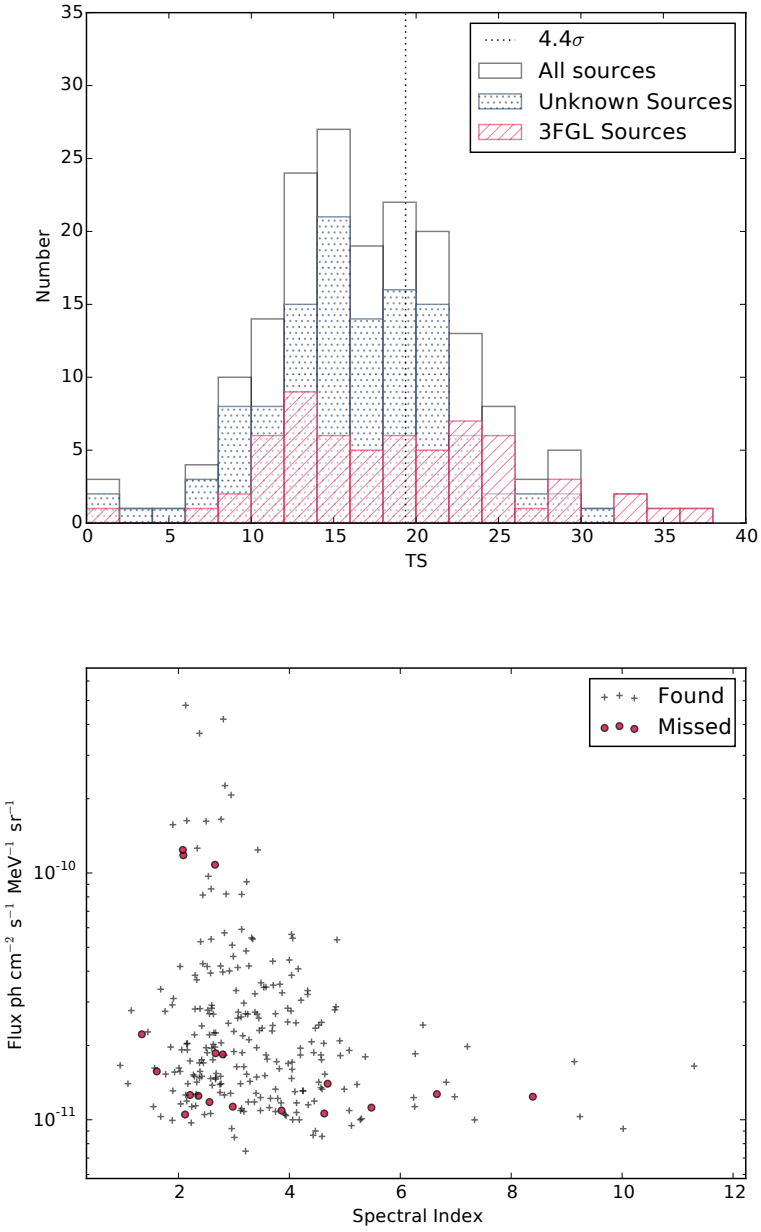


Figure 7.12: Top: The 50 GeV $< E_\gamma <$ 2 TeV TS distribution of the 131 sources found that did not correspond to 2FHL source positions. Bottom: The published spectral index and flux distribution for all 2FHL sources, where the black crosses are sources found with the DBSCAN algorithm and the red circles are those that are missed.

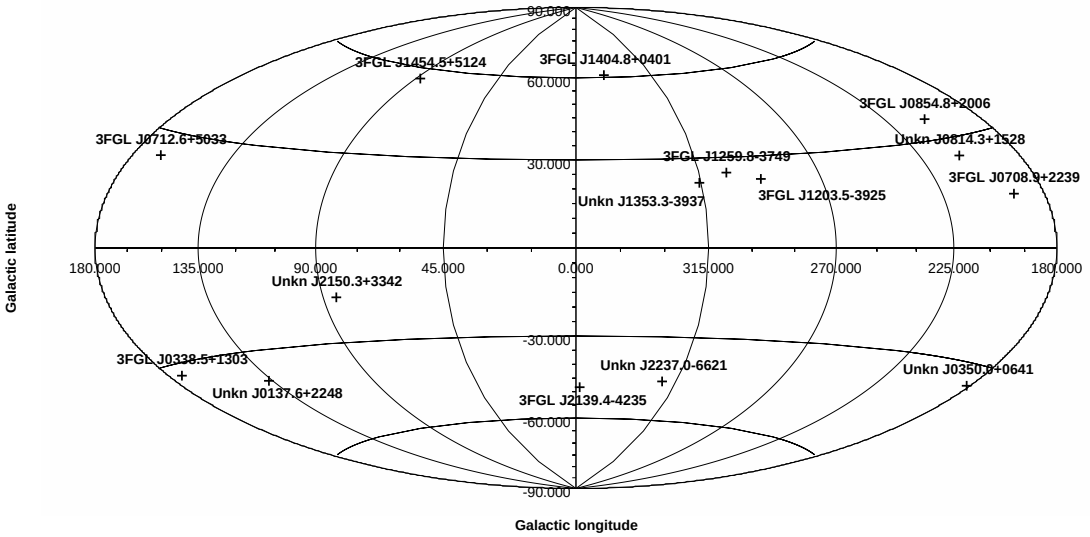


Figure 7.13: Clusters with TS >25 found between 50 GeV and 2 TeV using the same observation time as used to create the 2FHL.

2 of the Unkn. sources found in the 100 GeV to 3 TeV range are also found here, namely Unkn J0350.0+0641 and Unkn J0814.3+1528 where the former was seen to be spatially coincident with the BL Lac 2MASS J03495785+064126.

For the 20 2FHL sources that where not found in the 422 cluster sample, only 8 were not picked up at all by DBSCAN, with the rest falling below the probability cut. In Figure 7.12 the published flux and spectral index values are shown for all extragalactic 2FHL sources separated into those found by DBSCAN and those missed (both below the probability threshold and those missed altogether). As can be seen most sources are at the low end of the flux range, save for 3 which have large fluxes. Upon further investigation at least 2 of these, 2FHL J1221.3+3009 (PG 1218+304) and 2FHL J1217.9+3006 (1ES 1215+303) have been absorbed into the TeVCat source 1ES 1218+304 and the third is the extended source 2FHL J0526.6-6825e otherwise known as the LMC. This indicates that DBSCAN may not currently be optimally set up for preventing source confusion or detecting extended sources.

	Name	Ra [deg] J2000	Dec [deg] J2000	Association	Type	z	Var	TS	Flux $\times 10^{-11}$ ph cm^{-2} s^{-1}
1	Unkn J0137.6+2248	24.49	22.81	...	...	...	...	26.54	1.18 ± 0.69
2	3FGL J0338.5+1303	54.67	13.05	RX J0338.4+1302	BCU	...	58.88	32.89	1.62 ± 0.82
3	Unkn J0350.0+0641	57.53	6.60	...	...	...	...	30.29	1.43 ± 0.78
4	3FGL J0708.9+2239	107.22	22.78	GB6 J0708+2241	BCU	...	56.60	36.93	1.71 ± 0.82
5	3FGL J0712.6+5033	108.32	50.58	GB6 J0712+5033	BLL	...	94.44	28.38	1.14 ± 0.60
6	Unkn J0814.3+1528	123.51	15.49	...	...	...	...	25.75	1.04 ± 0.68
7	3FGL J0854.8+2006	133.57	20.15	OJ 287	BLL	0.306	1059.57	33.50	2.01 ± 0.88
8	3FGL J1203.5-3925	180.70	-39.46	PMN J1203-3926	BCU	...	40.03	25.69	1.57 ± 0.80
9	3FGL J1259.8-3749	195.04	-37.94	NVSS J125949-374856	BCU	...	44.91	35.33	1.42 ± 0.72
10	Unkn J1353.3-3937	208.36	-39.63	...	...	...	...	26.75	0.89 ± 0.55
11	3FGL J1404.8+0401	211.21	3.98	MS 1402.3+0416	BLL	0.344	40.29	25.48	1.06 ± 0.64
12	3FGL J1454.5+5124	223.66	51.40	TXS 1452+516	BLL	1.083	108.12	28.83	0.89 ± 0.52
13	3FGL J2139.4-4235	324.86	-42.61	MH 2136-428	BLL	...	241.08	28.40	1.05 ± 0.79
14	Unkn J2150.3+3342	327.62	33.70	...	...	...	...	29.55	1.31 ± 0.72
15	Unkn J2237.0-6621	339.26	-66.35	...	...	...	...	28.49	0.85 ± 0.53

Table 7.3: Clusters found with TS >25 found between 50 GeV and 2 TeV using the same observation time as used to create the 2FHL. Including known published associated object, type, redshift (z) and variability index (Var).

7.7.2 Post 2FHL

As a next step, the possibility of using DBSCAN as a tool to scan the entire sky over regular time periods is considered. In order to investigate this, a further 6 months of data since the release of the 2FHL are included and the same clustering and maximum likelihood method as described in Section 7.7 is applied in order to search for sources that have increased in significance over time. This increase in time results in a 4.5% increase in the number of events in the analysed data set (around 2800 events including the Galactic plane). This increase in statistics led to the detection of a further 19 sources that are not published within the 2FHL. These can be found in Table 7.4 and their positions seen in Figure 7.14. The new sources consist of 13 3FGL sources and 6 Unkn. Given these findings, it is evident that the DBSCAN method can be applied at regular intervals in order to produce a continuous catalogue of sources that have become significant within the high energy regime of *Fermi*-LAT. A future implementation of the analysis code which automatically searches for new sources is envisaged.

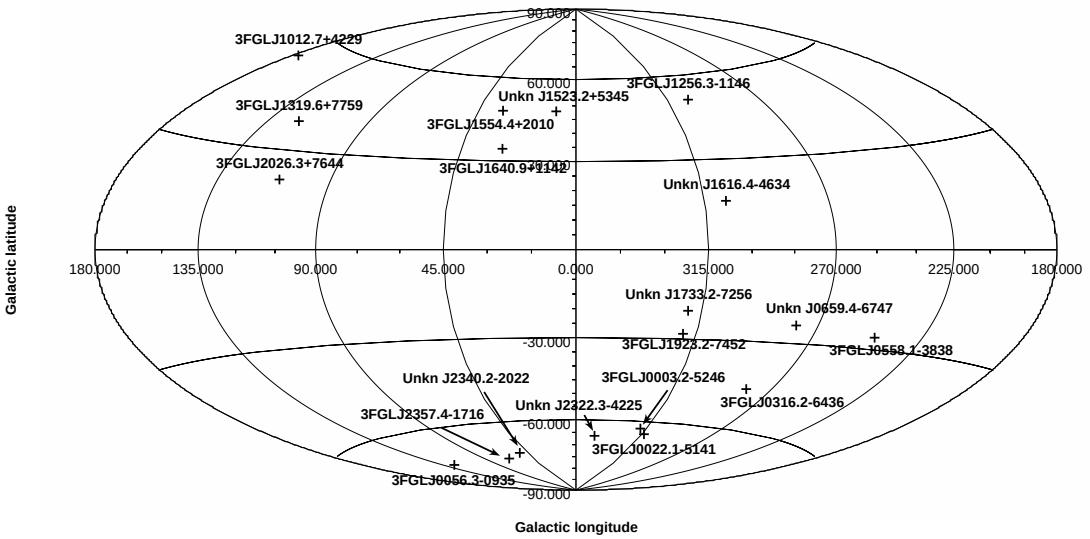


Figure 7.14: Clusters found with TS >25 found between 50 GeV and 2 TeV using the same observation time as used to create the 2FHL with an additional 6 months.

	Name	Ra [deg] J2000	Dec [deg] J2000	Association	Type	z	Var	TS	Flux $\times 10^{-11}$ ph cm^{-2} s^{-1}
1	3FGL J0003.2-5246	0.79	-52.80	RBS 0006	BCU	...	45.28	27.26	0.86 ± 0.50
2	3FGL J0022.1-5141	5.45	-51.65	1RXS J002159.2-514028	BLL	...	59.41	32.58	0.85 ± 0.50
3	3FGL J0056.3-0935	14.09	-9.65	TXS 0053-098	BLL	0.10312	55.52	26.52	1.91 ± 0.88
4	3FGL J0316.2-6436	49.09	-64.67	SUMSS J031614-643732	BLL	...	56.85	31.11	1.08 ± 0.55
5	3FGL J0558.1-3838	89.50	-38.67	EXO 0556.4-3838	BLL	0.302	37.60	32.34	1.18 ± 0.62
6	Unkn J0659.4-6747	104.90	-67.79			...	...	30.93	0.96 ± 0.52
7	3FGL J1012.7+4229	153.20	42.38	B3 1009+427	BLL	0.36513	50.43	25.65	1.15 ± 0.64
8	3FGL J1256.3-1146	194.08	-11.76	PMN J1256-1146	BCU	0.05791	59.44	30.46	1.06 ± 0.61
9	Unkn J1316.4-4634	199.15	-46.58			...	...	26.20	0.89 ± 0.53
10	3FGL J1319.6+7759	199.87	78.10	NVSS J131921+775823	BCU	...	51.89	29.41	1.04 ± 0.49
11	Unkn J1523.2+5345	230.85	5.58			...	...	25.26	0.92 ± 0.55
12	3FGL J1554.4+2010	238.58	20.17	1ES 1552+203	BLL	0.22229	31.45	35.94	1.18 ± 0.59
13	3FGL J1640.9+1142	250.20	11.69	TXS 1638+118	BLL	...	36.61	28.74	0.83 ± 0.50
14	Unkn J1733.2-7256	263.35	-72.94			...	...	29.20	0.89 ± 0.50
15	3FGL J1923.2-7452	291.05	-74.82		Unk	...	50.83	30.22	1.07 ± 0.54
16	3FGL J2026.3+7644	306.85	76.79	1RXS J202633.4+764432	BCU	...	31.96	25.80	0.73 ± 0.41
17	Unkn J2322.3-4225	350.65	-42.42			...	...	29.42	0.92 ± 0.54
18	Unkn J2340.2-2022	355.09	-20.38			...	...	25.57	1.88 ± 0.92
19	3FGL J2357.4-1716	359.37	-17.31	RBS 2066	BLL	...	44.08	32.16	1.29 ± 0.70

Table 7.4: Clusters found with TS >25 found between 50 GeV and 2 TeV using the same observation time as used to create the 2FHL with an additional 6 months. Including known published associated object, type, redshift (z) and variability index (Var).

7.8 Discussion and Conclusion

In the previous Chapter and in work presented in [24], the potential of using the clustering algorithm DBSCAN for performing an unbiased search for new sources within the extragalactic *Fermi*-LAT data was shown, detecting 9 new VHE AGN. Re-applying this to the improved Pass 8 events along with an additional 8 months of data, a further 36 VHE sources were found. These consist of 23 AGN and 12 previously undetected sources. Furthermore, in a comparison with the 2FHL catalogue of sources between 50 GeV and 2 TeV, 15 sources were discovered that were not presented in the 2FHL along with a further 19 using an additional 6 months of data.

In order to fully utilise the Pass 8 data set, it was necessary to extrapolate the Galactic and extragalactic background models up to 3 TeV. As the *Fermi*-LAT collaboration advise against using data above 500 GeV, results for analysis extending only up to this energy were also presented. In the data set used in this analysis, it is worth noting that the data above 500 GeV provide an additional 1138 events and have added to the significance of some of the sources presented in Table 7.1. In total, 10 sources can be seen to increase in significance with the increased energy range, 5 of which increase by at least a TS of 10. All but 1 of these are from the Unkn classification and are either insignificant or have a low significance in the $<$100 GeV range. This therefore makes them interesting targets for follow up observations.

As mentioned before, one of the reasons for applying this algorithm to VHE data was to increase the possible source list that could be followed up with ground based IACTs, where the energy threshold for observation is usually around 100 GeV. However, due to the pointing nature of these instruments and relatively short observation periods, it is important to know whether or not these sources are highly variable before follow up observations are considered. Therefore, for the sources detected that are also in the 3FGL, the variability index determined by the *Fermi*-LAT collaboration is shown, noting that only 7 out of a possible 24 are classed as variable. For the previously undetected objects, a search for any temporal clustering in the arrival time of each VHE event is performed, but no evidence of any clear grouping is seen.

Given the number of sources presented in this work, a total of 70 from both the 100 GeV to 3 TeV analysis and the comparison to the 2FHL, it is clear that the DBSCAN algorithm is a proficient tool at identifying sources within the *Fermi*-LAT data set. It is hoped that continued work will provide new sources during future analysis. A number of improvements are also planned, for example introducing a third dimension, such as event arrival time, to the clustering algorithm which will facilitate the search for more complex features hidden within the data.

Chapter 8

Summary, Conclusion and Future Work

The research presented in this thesis has focused on the simulated performance of GCT and working on expanding the known VHE AGN population that could be observed with CTA, using the *Fermi*-LAT data set as a pathfinder. In Chapter 2 a brief overview of the high energy universe was given followed by the important physical interactions that lead to the production and attenuation of gamma-rays. From there a discussion was given concerning the main source class of interest in the later chapters of this work, VHE AGN and more specifically the blazars which continue to be observable at GeV to TeV energies.

In Chapter 3 the instrumental techniques used to observe the HE to VHE sky were discussed. The current space based instrument *Fermi*, by way of its Large Area Telescope (LAT), has provided the community with the most detailed and indispens-

able full sky data set from energies of a few MeV up to 3 TeV to date. However, due to the small size of the LAT instrument, it is statistically limited at the higher energies. This is where ground-based gamma-ray instruments become important, making use of the cascades of particles produced by incoming VHE photons and the resulting Cherenkov radiation they are able to dramatically increase the effective area required to observe these events. The physics behind these airshowers and the technique used to observe them were described, which led to the introduction of the planned future ground based gamma-ray observatory, the Cherenkov Telescope Array (CTA), which aims to improve over current instruments by an order of magnitude. Finally, to observe the upper energy range that CTA intends to cover, it is foreseen that there will be a large component of the array that will be made up of many small size telescopes (SSTs). This leads to the description of the telescope that is the main focus of this work, one of the planned solutions to the SSTs, the Gamma Cherenkov Telescope (GCT).

With a prototype of GCT up and running in Paris, last November recording the first air shower image seen by a CTA prototype, there was a push to improve the model used within Monte Carlo simulations. This update for both a system using MaPMs (GCTM) and SiPMs (GCTS) was described in Chapter 4 which also acted to provide a better understanding of the inner working of the GCT telescope. Having completed this, the performance of the telescope was evaluated by showing that it could achieve a trigger efficiency at the required level (50% at 100 p.e.) and accurately reconstruct the observed charge within the camera. Using the evaluated telescope models, high level performance indicators such as angular resolution, energy resolution and differential sensitivities where derived for a range of telescope configurations. Using 5 arrays each with 7 telescopes at increasing inter-telescope distance, the relative performance of a "mini-array" of telescopes was derived. This represents the initial telescopes placed at the CTA observatory and can be used to estimate what would be achievable in the very early stages of construction.

As CTA will consist of many telescopes when complete, the majority of observed events are expected to fall within the array. Therefore in order to obtain a set of performance curves which better represent those that would be achievable by

the full SST component of CTA, a distance cut was applied to the "mini-array". This revealed that the expected angular and energy resolutions of the final CTA array above 1 TeV should easily be met by the the GCT contribution. Finally, by extrapolating the 7 telescope array to an array of 73 telescopes, it was shown that this is also true for the differential sensitivity. In all of the above, a general trend was seen for an improvement at high energies using a compact array in terms of energy and angular resolution and a more expanded array for sensitivity, although a full understanding of the optimised array layout requires further work. Finally, in terms of relative performance, it is seen that GCTS will outperform GCTM, but with further optimisation the two designs might become comparable.

In order to maximise the potential performance it is important that the telescope system be fully understood, which is achieved through a variety of calibration methods. It was briefly described in Chapter 3 that the camera for GCT will have 4 LED flashers, capable of producing a pulse of blue light that can be used to flatfield the camera. Another method, which has been used for current IACTs, is the use of Cherenkov radiation produced by local, unaccompanied muons. As there was an initial concern that, due to the small mirror area of the SSTs, insufficient muons would be detected to enable telescope calibration, an initial study investigating the feasibility of using muons for absolute calibration of GCT was presented in Chapter 5. As both the muon spectrum and the physics by which the muons produce Cherenkov radiation are well understood, the amount of light that should be observed for each muon event can be accurately predicted. Therefore, by measuring this amount of light over long periods of time, any loss in efficiency of the telescope system can be characterised by the parametrisation of the muon ring imaged within the camera. Using the Taubin method to fit muon rings, and theory developed in previous experiments, it was shown that although there is some difficulty created by the presence of a secondary mirror, both GCTM and GCTS should be able to use the muon calibration method. Improvements in the current method and an optimisation of the selection criteria and cuts will be required to fully understand the potential of the method.

While CTA will operate as an open observatory, there will be a large amount

of time dedicated to a set of key science projects. These will tackle some of the larger questions that CTA has set out to answer and require a sizeable amount of observation hours. It is foreseen that there will be deep observations of a selection of AGN to obtain high quality spectra along with long term monitoring. Additionally there will be an extragalactic survey which will identify many new sources. With this in mind, the later work in this thesis, presented in Chapter 6 and 7, focuses on using the clustering algorithm DBSCAN to search for potential VHE AGN within the *Fermi*-LAT data set. In Chapter 6 an initial study was carried out on Pass 7 reprocessed data in the energy range of 100-300 GeV for Galactic latitudes $|b| > 10°$. This revealed the presence of 9 previously unknown AGN that were significant in the VHE domain. In Chapter 7 a more sophisticated code and analysis was applied on the new Pass 8 *Fermi*-LAT data set which provided an extended energy range up to a possible 3 TeV. This work revealed a further 36 sources significant above 100 GeV and in a comparison with the 2FHL (80 months, 50 GeV - 2 TeV) an additional 15 sources were found. Using an extra 6 months from the 2FHL data set a further 19 sources were detected. From the discovery of these 70 sources, it is concluded that DBSCAN is a powerful tool for identifying sources within large gamma-ray data sets, helping to provide an unbiased catalogue of sources from the *Fermi*-LAT data set that could make promising targets for follow-up observations of current IACTs and CTA in the future. In addition, the identification of a larger data set of VHE AGN strengthens the justification for the CTA extragalactic survey which will reach much deeper sensitivities than the *Fermi* VHE data set.

Bibliography

[1] Aartsen, M. G., Abbasi, R., Abdou, Y., et al. 2013, *Search for Galactic PeV gamma rays with the IceCube Neutrino Observatory*, Physical Review D , 87, 062002

[2] Abramowski, A., Aharonian, F., Ait Benkhali, F., et al. 2014, *Diffuse Galactic gamma-ray emission with H.E.S.S.*, Physical Review D , 90, 122007

[3] Acharya, B. S., Actis, M., Aghajani, T., et al. 2013, *Introducing the CTA concept*, Astroparticle Physics, 43, 3

[4] Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, *Fermi Large Area Telescope First Source Catalog*, The Astrophysical Journal Supplement , 188, 405

[5] Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, *Spectrum of the Isotropic Diffuse Gamma-Ray Emission Derived from First-Year Fermi Large Area Telescope Data*, Physical Review Letters, 104, 10, 101101

[6] Ackermann, M., Ajello, M., Albert, A., et al. 2012, *Fermi LAT search for dark matter in gamma-ray lines and the inclusive photon spectrum*, Physical Review D , 86, 022002

[7] Ackermann, M., Ajello, M., Albert, A., et al. 2012, *The Fermi Large Area Telescope on Orbit: Event Classification, Instrument Response Functions, and Calibration*, The Astrophysical Journal Supplement , 203, 4

[8] Ackermann, M., Ajello, M., Albert, A., et al. 2012, *The Fermi Large Area Telescope On Orbit: Event Classification, Instrument Response Functions, and Calibration*, The Astrophysical Journal Supplement, 203, 4

[9] Ackermann, M., Ajello, M., Allafort, A., Antolini, E., Atwood, W. B., Axelsson, M., Baldini, L., Ballet, J. et al., 2012, *Fermi Large Area Telescope Second Source Catalog*, The Astrophysical Journal Supplement, 199, 31

[10] Acero, F., Ackermann, M., Ajello, M., et al. 2015, *Fermi Large Area Telescope Third Source Catalog*, The Astrophysical Journal Supplement , 218, 23

[11] Ackermann, M., Ajello, M., Atwood, W. B., et al. 2016, *2FHL: The Second Catalog of Hard Fermi-LAT Sources*, The Astrophysical Journal Supplement , 222, 5

[12] Actis, M., Agnetta, G., Aharonian, F., Akhperjanian, A., Aleksic, J. et al., 2011, *Design Concepts for the Cherenkov Telescope Array*, Exp. Astron., 32, 193

[13] Adrián-Martínez, S., Albert, A., et al. for the ANTARES Collaboration 2015, *ANTARES constrains a blazar origin of two IceCube PeV neutrino events*, Astronomy and Astrophysics, , 576, L8

[14] Aggarwal, C. C. and Reggy, C. K., 2014, *Data Clustering, Algorithms and Applications*, CRC Press, Chapman & Hall Book

[15] Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2006, *A low level of extragalactic background light as revealed by -rays from blazars*, Nature , 440, 1018

[16] Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2006, *Observations of the Crab nebula with HESS*, Astronomy and Astrophysics, 457, 899

[17] Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2007, *An Exceptional Very High Energy Gamma-Ray Flare of PKS 2155-304*, The Astrophysical Journal Letters , 664, L71

[18] Anderberg M. R., 1973, *Cluster Analysis For Applications*, Academic Press Inc, New York, Chapter 6.

[19] Anderhub, H., Backes, M., Biland, A., et al. 2013, *Design and operation of FACT - the first G-APD Cherenkov telescope*, Journal of Instrumentation, 8, P06008

[20] Ankerst, M., Markus M. Breunig, Hans-Peter Kriegel, and Jrg Sander, 1999, *OPTICS: ordering points to identify the clustering structure*, ACM Press, pp. 49-60

[21] Antoni, T., Apel, W. D., Badea, F., et al. 2001, *Electron, muon, and hadron lateral distributions measured in air showers by the KASCADE experiment* Astroparticle Physics, 14, 245

[22] Antonucci, R. 1993, *Unified models for active galactic nuclei and quasars*, Annual Rev. Astron. Astrophys. , 31, 473

[23] Armstrong, P., Chadwick, P. M., Cottle, P. J., et al. 1999, *The University of Durham Mark 6 Gamma Ray Telescope*, Experimental Astronomy, 9, 51

[24] Armstrong, T., Brown, A. M., Chadwick, P. M., & Nolan, S. J. 2015, *The detection of Fermi AGN above 100 GeV using clustering analysis*, MNRAS, 452, 3159

[25] Armstrong, T., Costantini, H., Rulten, C., et al. 2015, *Monte Carlo Studies of the GCT Telescope for the Cherenkov Telescope Array*, PoS(ICRC2015)942, arXiv:1508.06088

[26] Atwood, W.B. et al., 2009, *The Large Area Telescope on the Fermi Gamma-ray Space Telescope Mission*, The Astrophysical Journal, 697,1071

[27] Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, *The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission*, The Astrophysical Journal , 697, 1071

[28] Atwood, W., Albert, A., Baldini, L., et al. 2013, *Pass 8: Toward the Full Realization of the Fermi-LAT Scientific Potential*, arXiv:1303.3514

[29] Atwood, W., Albert, A., Baldini, L., et al. 2013, *Pass 8: Toward the Full Realization of the Fermi-LAT Scientific Potential*, arXiv:1303.3514

[30] Benn, C. R., & Ellison, S. L. 1998, *Brightness of the night sky over La Palma*, New Astronomy Reviews, Vol. 42, p. 503-507

[31] Berge, D, 2002, *Development of an Algorithm for the Shower Reconstruction with the H . E . S . S . Telescope System*, Masters Thesis, Berlin Humboldt University, http://berge.web.cern.ch/berge/Dipl_DB.pdf

[32] Bernlöhr, K. 2008, *Simulation of imaging atmospheric Cherenkov telescopes with CORSIKA and sim_ telarray*, Astroparticle Physics, 30, 149

[33] Bernlöhr, K. 2001, *CORSIKA and sim_telarray Simulation of the imaging atmospheric Cherenkov technique*. software documentation

[34] Blandford, R. D., Königl, A. 1979, *Relativistic jets as compact radio sources.* The Astrophysical Journal , 232, 34

[35] Blinov, D., Pavlidou, V., Papadakis, I., et al. 2015, *RoboPol: first season rotations of optical polarization plane in blazars*, Monthly Notices of the Royal Astronomical Society, 453, 1669

[36] Bolz, O. 2004 *Absolute Energiek alibration der abbildenden Cherenkov-Teleskope des H.E.S.S. Experiments und Ergebnisse erster Beobachtungen des Supernoa-Uberrests RX J1713.7-3946*, PhD Thesis, Universitat Heidelberg

[37] Böttcher, M., Reimer, A., Sweeney, K., & Prakash, A. 2013, *Leptonic and Hadronic Modeling of Fermi-detected Blazars*, The Astrophysical Journal 768, 54

[38] Böttcher, M., Reimer, A., & Marscher, A. P. 2009, *Implications of the very High Energy Gamma-Ray Detection of the Quasar 3C279*, The Astrophysical Journal 703, 1168

[39] Bouvier, A., Gebremedhin, L., Johnson, C., et al. 2013, *Photosensor characterization for the Cherenkov Telescope Array: silicon photomultiplier ver-*

sus multi-anode photomultiplier tube Proceedings of the SPIE, Vol. 8852, ID 88520K

[40] Bugaev, E., Misaki, A., Naumov, V., Sinegovskaya, T., Sinegovsky, S. and Takahashi, N., 1998, *Atmospheric muon flux at sea level, underground, and underwater*, Physical Review D, Vol 58(5) pp. 054001

[41] Bregeon, J., Charles, E., & M. Wood for the Fermi-LAT collaboration 2013, *Fermi-LAT data reprocessed with updated calibration constants*, arXiv:1304.5456

[42] Brown, A. M., Armstrong, T., Chadwick, P. M., et al. 2015, *Flasher and muon-based calibration of the GCT telescopes proposed for the Cherenkov Telescope Array*, arXiv:1509.00185

[43] Brown A.M., 2014, *Very high energy γ-ray emission from RBS 0970* MNRAS, 442, L56-L60

[44] Brown A.M., Chadwick P.M., Landt, H., 2014, *Very high energy γ-ray emission from RBS 0679*, MNRAS, 445, 4345-4350

[45] Calore, F., Cholis, I., & Weniger, C. 2015, *Background model systematics for the Fermi GeV excess*, Journal of Cosmology and Astroparticle Physics, 3, 038

[46] Campana, R., Massaro, E., Gasparrini, D., Cutini, S., & Tramacere, A. 2008, *A MST algorithm for source detection in gamma-ray images*, MNRAS, 383, 1166

[47] Carlson E., Linden T., Profumo S., and Weniger C., 2013, *Clustering analysis of the morphology of the 130 GeV gamma-ray feature*, Physical Review D, 88, 043006

[48] Casandjian, J.-M., Grenier, I., & for the Fermi Large Area Telescope Collaboration 2009, *High Energy Gamma-Ray Emission from the Loop I region*, arXiv:0912.3478

[49] Chalme-Calvet, R., de Naurois, M., Tavernet, J.-P., & for the H. E. S. S. Collaboration, 2014, *Muon efficiency of the H.E.S.S. telescope*, arXiv:1403.4550

[50] Ciprini, S., Tosti, G., Marcucci, F., et al. 2007, *1D, 2D, 3D wavelet methods for gamma-ray source analysis*, The First GLAST Symposium, 921, 546

[51] CLAS12 RICH Collaboration, 2013, *CLAS12 RICH Conceptual Design Report*, Unpublished manuscript.

[52] Clay, J/, 1927, *Penetrating Radiation I.* Proc. Amsterdam Acad.Sci., 30, 1115

[53] Cottingham, W. M. and Greenwood D. A., 2007 *An Introduction to the Standard Model of Particle Physics* Cambridge Publishing

[54] Damiani, F., Maggio, A., Micela, G., & Sciortino, S. 1997, *A Method Based on Wavelet Transforms for Source Detection in Photon-counting Detector Images. II. Application to ROSAT PSPC Images*, APJ, 483, 350

[55] Daum, A., Hermann, G., Heß, M., et al. 1997, *First results on the performance of the HEGRA IACT array*, Astroparticle Physics, 8, 1

[56] De Franco, A., White, R., Allan, D., et al. 2015, *The first GCT camera for the Cherenkov Telescope Array*, arXiv:1509.01480

[57] Deligny, O., de Mello Neto, J., Sommers, P., et al. 2013, *Review of the anisotropy working group at UHECR-2012*, European Physical Journal Web of Conferences, 53, 01008

[58] Dobler, G., Finkbeiner, D. P., Cholis, I., Slatyer, T., & Weiner, N. 2010, *The Fermi Haze: A Gamma-ray Counterpart to the Microwave Haze*, The Astrophysical Journal , 717, 825

[59] Dournaux, J.-L., Huet, J.-M., Dumas, D., et al. 2015, *Performance of the Mechanical Structure of the SST-2M GCT Proposed for the Cherenkov Telescope Array*, arXiv:1508.06415

[60] Dournaux, J.-L., Huet, J.-M., Amans, J.-P., et al. 2014, *Mechanical Design of SST-GATE, a Dual-Mirror telescope for the Cherenkov Telescope Array*, Proc. SPIE, 9151, 915102

[61] Dubus, G., Contreras, J. L., Funk, S., et al. 2013, *Surveys with the Cherenkov Telescope Array*, Astroparticle Physics, 43, 317

[62] Dumas, D., Laporte, P., Sol, H., et al. 2014, *SST dual-mirror telescopes for the Cherenkov Telescope Array*, Proc. SPIE, 9145, 91452Z

[63] Dumas, D., Huet, J.-M., Dournaux, J.-L., et al. 2014, *SST dual-mirror telescope for Cherenkov Telescope Array: an innovative mirror manufacturing process*, Proc. SPIE, 9151, 91512U

[64] Earnshaw, J. C., 1968, *The Muon Component of Extensive Air Showers*, Durham Thesis, Durham University

[65] Ester M., Kriegel H.P., Sander J., and Xu X., 1996, *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*, AAAI Press, pp. 226-231.

[66] Fegan, D. J. 1997, *γ/hadron separation at TeV energies*, Journal of Physics G Nuclear Physics, 23, 1013

[67] Fleury P., and Artemis-Whipple Collaboration, 1991, *Čerenkov ring images of cosmic ray muons.*, International Cosmic Ray Conference, Vol 2. p595-598

[68] Gaidos, J. A., Akerlof, C. W., Biller, S., et al. 1996, *Extremely rapid bursts of TeV photons from the active galaxy Markarian 421*, Nature , 383, 319

[69] Gaug, M., Armstrong, T., Bernlohr, K., Daniel, M., Errando, M., Maccarrone, M.C., Majumdar, P., Mineo, T., Mitchell, A., Moderski, R., Parsons, D., Prandini, E. and Toscano, E. *Using Muon Rings for the Optical Throughput Calibration of the Cherenkov Telescope Array*, CTA internal document.

[70] F.Goebel , K. Mase , M. Meyer, R. Mirzoyan , M. Shayduk , and M. Teshima, 2005, *Absolute energy scale calibration of the MAGIC telescope using muon images*, 29th International Cosmic Ray Conference, 00, 101106

[71] Hartigan, J. A., 1975, *Clustering Algorithms*, John Wiley & Sons, Inc. Chapter 4,

[72] Healey S. E., Romani R. W., Taylor G. B., Sadler E. M., Ricci R., Murphy T., Ulvestad J. S., and Winn J. N., 2007, *CRATES: An All-Sky Survey of Flat-Spectrum Radio Sources*, ApJS, 171, , pp. 61-71.

[73] Hebbeker, T. and Timmermans, C., 2002, *A compilation of high energy atmospheric muon data at sea level*, Astroparticle Physics, Vol. 18, pp. 107-127

[74] Heck, D., Knapp, J. Capdevielle, J. N. Schatz, G. and Thouw, T., 1998, *CORSIKA: a Monte Carlo code to simulate extensive air showers.*, Forschungszentrum Karlsruhe GmbH, Karlsruhe (Germany), V + 90 p., TIB Hannover, D-30167 Hannover (Germany).

[75] Hess, V. F., 1912, *Durchdringende Strahlung Bei Sieben Freiballonfahrten*, Physik, Zwitschr. XIII, 1084

[76] Hillas, A. M. 1985, *Cherenkov light images of EAS produced by primary gamma rays and by nuclei*, International Cosmic Ray Conference, 3,

[77] Hinton, J. A., & Hofmann, W. 2009, *Teraelectronvolt Astronomy* Annual Review of Astronomy & Astrophysics, 47, 523

[78] Hinton, J. A., & the HESS Collaboration 2004, *The status of the HESS project*, New Astronomy Reviews, 48, 331

[79] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS A CAT 040 (2007) [physics/0703039].

[80] Kampert, K.-H. 2007, *Cosmic Rays from the Knee to the Ankle - Status and Prospects* - Nuclear Physics B Proceedings Supplements, 165, 294

[81] Kerr, M. 2010, *Likelihood Methods for the Detection and Characterization of Gamma-ray Pulsars with the Fermi Large Area Telescope*, PhD thesis, University of Washington, ArXiv:1101:6072

[82] Kharb, P., Lister, M. L., & Cooper, N. J. 2010, *Extended Radio Emission in MOJAVE Blazars: Challenges to Unification*, The Astrophysical Journal , 710, 764

[83] Kildea, J., Atkins, R. W., Badran, H. M., et al. 2007, *The Whipple Observatory 10 m gamma-ray telescope, 1997 2006*, Astroparticle Physics, 28, 182

[84] Leroy, N., Bolz, O., Guy, J., et al. 2003, *Calibration Results for the First Two H.E.S.S. Array Telescopes*, International Cosmic Ray Conference, 5, 2895

[85] Li, T. P. and Ma, T. Q., 1983, *Analysis methods for results in gamma-ray astronomy*, Astrophysical Journal, 272, 317

[86] Lorenz, E., & Martinez, M. 2005, *High energy astrophysics: The MAGIC telescope*, Astronomy and Geophysics, 46, 6.21

[87] Martin, B. R. and Shaw, G., 1999 *Particle Physics* Wiley Publishing, Second Edition

[88] Massaro, E., Giommi, P., Leto, C., et al. 2009, *Roma-BZCAT: a multifrequency catalogue of blazars*, AAP, 495, 691

[89] Matthews, J. 2005, *A Heitler model of extensive air showers*, Astroparticle Physics, 22, 387

[90] Mazin, D. 2009, *Constraints on Extragalactic Background Light from Cherenkov telescopes: status and perspectives for the next 5 years*, American Institute of Physics Conference Series, 1112, 111

[91] Meegan, C., Lichti, G., Bhat, P. N., et al. 2009, *The Fermi Gamma-ray Burst Monitor*, The Astrophysical Journal , 702, 791-804

[92] Millikan, R. A. and Cameron, G. H., 1928 *The Origin of the Cosmic Rays*, Phys. Rev., 42, 4, pp 533-557

[93] Mori, M., & CANGAROO Collaboration 2001, *The CANGAROO-III Project: Status report*, International Cosmic Ray Conference, 7, 2831

[94] Nagano, M., & Watson, A. A. 2000, *Observations and implications of the ultrahigh-energy cosmic rays*, Reviews of Modern Physics, 72, 689

[95] Neddermeyer S. H. and Anderson C. D. 1937, *Note on the Nature of Cosmic-Ray Particles*, Physical Review Vol. 51

[96] Niemiec, J., Michałowski, J., Dyrda, M., et al. 2013, *Single-Mirror Small-Size Telescope structure for the Cherenkov Telescope Array*, arXiv:1307.4189

[97] Okumura, A. 2011, *Development of Non-sequential Ray-tracing Software for Cosmic-ray Telescopes*, International Cosmic Ray Conference, 9, 211, astro-ph/1110.4448

[98] Okumura, A., Noda, K., & Rulten, C. 2016, *ROBAST: Development of a ROOT-based ray-tracing library for cosmic-ray telescopes and its applications in the Cherenkov Telescope Array*, Astroparticle Physics, 76, 38

[99] Olive, K. A., & Particle Data Group, 2014, *Review of Particle Physics*, Chinese Physics C, 38, 090001

[100] Pareschi, G., Agnetta, G., Antonelli, L. A., et al. 2013, *The dual-mirror Small Size Telescope for the Cherenkov Telescope Array*, arXiv:1307.4962

[101] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Journal of Machine Learning Research 12, 2825 (2011). *Scikit-learn: Machine Learning in Python*

[102] Pierog, T. 2013, *LHC data and extensive air showers*, European Physical Journal Web of Conferences, 52, 03001

[103] Prantzos, N., Boehm, C., Bykov, A. M., et al. 2011, *The 511 keV emission from positron annihilation in the Galaxy*, Reviews of Modern Physics, 83, 1001

[104] Roy S., and Bhattacharyya D. K., 2005, *An Approach to Finding Embedded Clusters Using Density Based Techniques*, Springer, pp. 523-535

[105] Rulten, C. B, 2012, *Performance studies for the Cherenkov Telescope Array (CTA) with prospects for detecting pulsed gamma-ray emission*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/3632/

[106] Rulten, C. B., Zech, A., Okumura, A. Laporte, P. Schmoll, J., 2016, *Simulating the optical performance of a small-sized telescope with secondary optics for the Cherenkov Telescope Array*, Astroparticle Physics, doi:10.1016/j.astropartphys.2016.05.002

[107] Schmelling, M., Hashim, N., Grupen, C., Luitz, S., Maciuc, F., Mailov, A., Mller, A., Sander, H., Schmeling, S., Tcaciuc, R., Wachsmuth, H., Ziegler, T. and Zuber K, 2013, *Spectrum and charge ratio of vertical cosmic ray muons up to momenta of 2.5 TeV/c*, Astroparticle Physics, Vol. 49, pp. 1-5

[108] Schönfelder, V. 2001, *The Universe in Gamma Rays*, 407

[109] Serrano, J., Alvarez, P., et al. 2009, *The white rabbit project*, in Proceedings of ICALEPCS TUC004

[110] Sokolov, A., & Marscher, A. P. 2005, *External Compton Radiation from Rapid Nonthermal Flares in Blazars*, The Astrophysical Journal 629, 52

[111] Sol, H., Zech, A., Boisson, C., et al. 2013, *Active Galactic Nuclei under the scrutiny of CTA*, Astroparticle Physics, 43, 215

[112] Sol, H., Zech, A., Boisson, C., et al. 2013, *Active Galactic Nuclei under the Scrutiny of CTA*, Astroparticle Physics, 43, 215

[113] Starck J. L. and Pierre M., 1998, *Structure Detection in Low Intensity X-Ray Images*, AAPS, 128, 397-407

[114] Taubin, G. 1991 *Estimation Of Planar Curves, Surfaces And Nonplanar Space Curves Defined By Implicit Equations, With Applications To Edge And Range Image Segmentation* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 13, pp. 1115-1138

[115] Tibaldo, L., Vandenbroucke, J. A., Albert, A. M., et al. 2015, *TARGET: toward a solution for the readout electronics of the Cherenkov Telescope Array*, arXiv:1508.06296

[116] Tramacere A., and Vecchio C., 2013, *γ-ray DBSCAN: a clustering algorithm applied to Fermi-LAT γ-ray data*, A&A, 549, A138

[117] Urry, C. M., & Padovani, P. 1995, *Unified Schemes for Radio-Loud Active Galactic Nuclei*, Publications of the Astronomical Society of the Pacific, 107, 803

[118] Vacanti, G., Fleury, P., Jiang, Y., Paré, E., Rovero, A. C., Sarazin, X., Urban, M. and Weekes, T. C. 1994, *Muon ring images with an atmospheric Čerenkov telescope*, Astroparticle Physics, Vol. 2, p1-11

[119] Vassiliev, V. V., & Fegan, S. J. 2008, *Schwarzschild-Couder two-mirror telescope for ground-based γ-ray astronomy*, International Cosmic Ray Conference, 3, 1445

[120] Véron-Cetty, M.-P., & Véron, P. 2010, *A catalogue of quasars and active nuclei: 13th edition* A&A, 518, A10

[121] Weekes, T. C., Badran, H., Biller, S. D., et al. 2002, *VERITAS: the Very Energetic Radiation Imaging Telescope Array System*, Astroparticle Physics, 17, 221

[122] Weeks, T. C., 2003 *Verhy High Energy Gamma-Ray Astronomy* IoP Publishing

[123] White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, *A Catalog of 1.4 GHz Radio Sources from the FIRST Survey*, ApJ, 475, 479

[124] York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *The Sloan Digital Sky Survey: Technical Summary*, The Astronomical Journal, 120, 1579

[125] Yusef-Zadeh, F., Hewitt, J. W., Wardle, M., et al. 2013, *Interacting Cosmic Rays with Molecular Clouds: A Bremsstrahlung Origin of Diffuse High-energy*

Emission from the Inner $2^\circ \times 1^\circ$ *of the Galactic Center* ,The Astrophysical Journal , 762, 33

[126] Zech, A., Amans, J.-P., Blake, S., et al. 2013, *SST-GATE: A dual mirror telescope for the Cherenkov Telescope Array* arXiv:1307.3035

[127] Zhong, S., Gosh, J.. 2003, *A Unified Framework for Model-Based Clustering*, Journal of Machine Learning Research, 4, 1001-1037

Appendix A

List of Acronyms and Abbreviations

2FGL - The second Fermi-LAT catalog

2FHL - The Second Catalog of Hard Fermi-LAT Sources

3FGL - The Third Fermi-LAT source catalog

AGN - Active Galactic Nuclei

BCU - Active Galactic Nuclei of Uncertain Type

BPWL - Broken Power Law

CHEC - Compact High Energy Camera

CTA - Cherenkov Telescope Array

DACQ - Data Acquisition system

DBSCAN - Density-Based Spatial Clustering of Applications with Noise

EAS - Extensive Air Showers

EBL - Extragalactic Background Light

EPWL - Power Law With Exponential cut off

FADC - Flash Analogue to Digital Converter

FOV - Field of View

FSRQ - Flat Spectrum Radio Quasar

GCT - Gamma Cherenkov Telescope

GCTM - GCT with camera using MaPMs

GCTS - GCT with camera using SiPMs

HV - High Voltage

IACT - Imaging Atmospheric Telescope

IRF - Instrument Response Function

LAT - Large Area Telescope

L/M/SST - Large/Medium/Small Size Telescope

LRT - Likelihood Ratio Test

MaPM - Multi-Anode Photomultiplier

MC - Monte Carlo

MET - Mission Elapsed Time

MSCRW - Mean Scaled Reduced Width

MSCRL - Mean Scaled Reduced Length

NSB - Night Sky Background

p.e. - photo electron

PDE - Photon Detection Efficiency

PSF - Point Spread Function

PWL - Power Law

SiPM - Silicon Photomultiplier

SPE - Single Photo Electron

SST-GATE - Small Size Telescope GAmma-ray Telescope Elements

TMVA - Toolkit for Multivariate Analysis

TS - Test Statistic

VHE - Very High Energy

Appendix B

Full List of sim_telarray Configuration Parameters

parameter	GCTM	GCTS
mirror class	2	
Mirror Shapes	asymetric (eqn. 4.6.1 & **??**)	
Focal Plane Shape	$\sim$ 1 m radius of curvature	
Focal length	2.228 m	
Primary Diameter	4 m (six petals)	
Secondary Diameter	2 m (monolithic)	
Camera diameter	42 cm	
Camera depth	50 cm	
random focal length	0.0	
Mirror Reflection Random Angle	0.0075	
mirror align random distance	0.0	
mirror align random horizontal	0.0	
mirror align random vertical	0.0	
mirror offset	0.0	
focus offset	0.0	
Primary Mirror Relectivity	Figure 4.5	
Secondary Mirror Reflectivity	Figure 4.5	
telescope random angle	0.0	
telescope random error	0.0	
Telescope Transmission	eqn. 4.6.2	
pixels parallel	0.0	
num gains	1	
Number of pixels	2048	
Quantum efficiency	Figure 4.10	
SPE Response	Figure 4.11	
PM voltage variation	0.03	0.0
PM transit time	5.3	4
transit time jitter	0.28	0.2
Gain variation	0.0	0.05

Table B.1: Table of Configuration file parameters for GCT-M and GCT-S

parameter	GCTM	GCTS
QE Variation	0.04	
NSB	0.0142	0.0410
Discriminator Bins	120	
Discriminator Start	3	
Discriminator Amplitude	20	
Discriminator Shape	Figure 4.12	
Discriminator Threshold	172	230
Trigger Pixels	2	
Discriminator Threshold variation	2	
discriminator time over threshold	1.0	
discriminator var time over threshold	0.0	
discriminator sigsum over threshold	0.0	
discriminator var sigsum over threshold	0.0	
discriminator hysteresis	0.0	
discriminator gate length	8.0	
discriminator var gate length	1	
discriminator output amplitude	42	
discriminator output var percent	0	
discriminator rise time	1.0	
discriminator fall time	1.0	
default trigger	Majority	
teltrig min time	0.5	
teltrig min sigsum	0.0	
trigger delay compensation	0,0,0	
fadc mhz	1000	
fadc pulse shape	Figure 4.12	
fadc bins	128	
fadc sum bins	96	
fadc sum offset	24	

Table B.2: Table of Configuration file parameters for GCT-M and GCT-S

parameter	GCTM GCTS
photon delay	5
fadc max signal	65535
fadc pedestal	40
fadc amplitude	8
fadc noise	2
pulse analysis	-9
sum before peak	7
sum after peak	10
tailcut scale	1.0

Table B.3: Table of Configuration file parameters for GCT-M and GCT-S